

Qy 61 ACTTCCTGTTGATTTAGTGGCAATTAGTAGTGAAACATCGAGACCCGACATTC 120
Db 61 ACTTCCTGTTGATTTAGTGGCAATTAGTAGTGAAACATCGAGACCCGACATTC 120
Qy 121 TATCCTGTGAAAGCAGGTTGTAACAGGCACAGTTGAAAGATTTGAGGGAGG 180
Db 121 TATCCTGTGAAAGCAGGTTGTAACAGGCACAGTTGAAAGATTTGAGGGAGG 180
Qy 181 CTATGATGTTTCAAGGCCAACATGCGGAAATGTCAGATCCTCTGAATATGA 240
Db 181 CTATGATGTTTCAAGGCCAACATGCGGAAATGTCAGATCCTCTGAATATGA 240
Qy 241 CATCAAACCTGGCATCAGATACTCGCGGTAAGACGGGCTTCTGTTAGCGGG 300
Db 241 CATCAAACCTGGCATCAGATACTCGCGGAAATGTCAGATCCTCTGAAGCGGG 300
Qy 301 TGACTGCGGAGCCAGGAACTGGGTGATGTTCTACTGGCGCTGAACTCAAGA 360
Db 301 TGACTGCGGAGCCAGGAACTGGGTGATGTTCTACTGGCGCTGAACTCAAGA 360
Qy 361 CGATAAATTCTGGCTCGGAATCGTGAGATGACTAACTACCAAMCGTCAACC 420
Db 361 CGATAAATTCTGGCTCGGAATCGTGAGATGACTAACTACCAAMCGTCAACC 420
Qy 421 GCTGAACCTGACGATTGTCAGCGAACCTACTGRCGCGAACATGTTGACTGT 480
Db 421 GCTGAACCTGACGATTGTCAGCGAACCTACTGRCGCGAACATGTTGACTGT 480
Qy 481 CAGCGAGAAAGCTCGTAGCGTTGAAAGCTCAACTCAATGGTTGGAAOTGACATAA 540
Db 481 CAGCGAGAAAGCTCGTAGCGTTGAAAGCTCAACTCAATGGTTGGAAOTGACATAA 540
Qy 541 CAGCCGTTCTGTTGATCGGCCCTGATGCGGAACTACAGTCGCGCGAGCTTCG 600
Db 541 CAGCCGTTCTGTTGATCGGCCCTGATGCGGAACTACAGTCGCGCGAGCTTCG 600
Qy 601 CGCCATGCTRACTCTCTGAAACCGAACATTAGTCGCGCGTGTAACTGCTGAA 660
Db 601 CGCCATGCTRACTCTCTGAAACCGAACATTAGTCGCGCGTGTAACTGCTGAA 660
Qy 661 GAAATGACGGCGCTGAAACCGAACATTAGTCGCGCGTGTAACTGCTGAA 660
Db 661 GAAATGACGGCGCTGAAACCGAACATTAGTCGCGCGTGTAACTGCTGAA 660
Qy 720 Continuation (47 of 47) of U00096 from base 460001 (U00096 Escherichia coli K-12 MG1655
Db 720 Query Match 100.0%; Score 927; DB 1; Length 39675;
Db 720 Best Local Similarity 100.0%; Pred. No. 2.3e-264; Matches 927; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 720 Qy 1 GTCATGTTACCGCGATCTGTAAATTGACATGCGCACATTACCG 780
Db 780 Qy 781 TAAGGTTATCGCTTCGGTATCGGAAATTATCGCTTACACCGCTAAAAA 840
Db 781 TAAGGTTATCGCTTCGGTATCGGAAATTATCGCTTACACCGCTAAAAA 840
Db 840 Qy 841 ACGGGCGCTTTAGCCCGTTTATTTCACTTATTCAGATACTGAACTCATC 900
Db 841 ACGGGCGCTTTAGCCCGTTTATTTCACTTATTCAGATACTGAACTCATC 900
Qy 900 Qy 901 GTCCGTGTTAATCTCTTACTGGCTT 927
Db 901 GTCCGTGTTAATCTCTTACTGGCTT 927
RESULT 2
U00096_46/c
WPCOMMENT
Sequence split into 47 fragments Locus U00096 Accession U00096
Fragment Name Begin End
U00096_00 1 110000
U00096_01 1 210000
U00096_02 200001 310000
U00096_03 300001 410000
U00096_04 400001 510000
U00096_05 500001 610000

CDS

complement (4205..4921)

/function="arcA"

/note="regulator; Global regulatory functions"

to residues 1 to 238 of 238 are 100.00 pct identical

strain MG1655; B4401"

/codon_start=1

/transl_table=11

/product="negative response regulator of genes in aerobic

pathways, (sensors, ArcB and CpxA)"

/protein_id="AGS5581_1"

/db_xref="GI:12519433"

/translation="MOPHILIVVDEPLVLENTLASIFEARQYVFTEATGAEHQIIS

EVDINLIVMDPGLPGANGLILAREQANVALFLTSGDNEVKLGIEGADDYT

KPFNPRLTRIRRNALISRTWMLGRIVTSEERSVSVEKFNGWELLINNSLIGPQEY

LFRSERFAMLAFFCECPGKQSRAEULKKMGTRELKPHRTDVTIRRHKFFSTPD

5017..5157

/gene="yjjY"

/note="synonym: 26005"

5017..5157

/gene="YJJY"

/function="orf; Unknown function"

/note="Residues 1 to 46 of 46 are 100.00 pct identical to

MG1655; B4402"

/codon_start=1

/transl_table=11

/product="orf, hypothetical protein"

/protein_id="AGS5582_1"

/db_xref="GI:12519434"

/translation="MTKVRNCVLDALSTINVNNIISLWVGTFPQDPVSKTAVILILT

AT"

/gene="yabT"

/note="Synonym: 26006"

5357..5623

/transl_table=11

/product="orf, hypothetical protein"

/protein_id="AGS5583_1"

/db_xref="GI:12519435"

/translation="MRITILIVVPARAENIGAACARAKMKGFSSTERIVSQAHPEPAT

RWWATSGSDIDNKRIFTLAEISLHDVFTVATRSRKYHAYTPVLPVLEPKS

SWSHAAVLFEREDSLTNEBLADLADLVGYPVMADYPSLXNGXAVXTCYQLALIQ

OPTKSQTADODHOQLQALERAMALLITLAVADDIKLVDLQORLGLIEORDTAMLHL

LHDIEKNITK"

ORIGIN

Query Match Best Local Similarity 99.3%; Score 920.6; DB 1; Length 6794; Matches 923; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GRCATGTTACCGCAGATGTTATTGCAAGCATGTCATCGAGGAGGTGAGGACTTCTT 60

Db 5021 GTCATGTTACCGCAGATGTTATTGCAAGCATGTCATCGAGGAGGTGAGGACTTCTT 60

Qy 61 ACTTCCGTTGATTGATTGTTGGCAATTAGTCAACATGAGACCCGCACATTCT 120

Db 4961 ACTTCCGTTGATTGATTGTTGGCAATTAGTCAACATGAGACCCGCACATTCT 120

Qy 121 TATGTTGAACGAGCTGGTAACAGGAACACGGTGAAGAATGATTGAGGGAGG 180

Db 4901 TATGTTGAACGAGCTGGTAACAGGAACACGGTGAAGAATGATTGAGGGAGG 180

Qy 181 TATGTTGAACGAGCTGGTAACAGGAACACGGTGAAGAATGATTGAGGGAGG 4842

Db 4841 TATGTTGAACGAGCTGGTAACAGGAACACGGTGAAGAATGATTGAGGGAGG 4782

CDS

241

CATCACCACTGGTATCATGATCATCTGCAGGGTAAAGAACGGTCTCTGCTTAGGGCG

300

4781 CATCACCACTGGTATCATGATCATCTGCAGGGTAAAGAACGGTCTCTGCTTAGGGCG

4722

301 TGAACCTGGCAGGCGCAATGTTGCTGATGTTCTGACTGGGGTGACAACCAAGT

360

4721 TGAACCTGGCAGGCGCAATGTTGCTGACTGGGGTGACAACCAAGT

4662

361 CGATAAACTCTGGCTCGAATCGGAGATGACTACATCACCACCGTACACC

420

4661 CGATAAACTCTGGCTCGAATCGGAGATGACTACATCACCACCGTACACC

4602

421 GCGTGAAGTGGCATGGACGAACTACTGTCCTGACCTGATGTTGGFACTGT

480

4601 GCCTGAGCTGAGATGCTGAGCAACCTGTCCTGACCTGATCCTGGFACTGT

4542

481 CAAGGAGAACCTGGCTGGCTGAAGACTAACGTCAGTGGACTGGACATCAA

540

4541 CAGCCGTTCTGTTGATGCTGGCCCTGATGGGAGCAGTACACAACTGCGGAGCTGGTCCG

600

601 CCCCATCTCTGCTGAACTGCTGAACTGCTGAACTGCTGAACTGCTGAA

660

4421 CGCCATCTCTGCTGAACTGCTGAACTGCTGAACTGCTGAACTGCTGAA

4362

461 GAAATGACGGCTGTGCTGAACTGCTGAACTGCTGAACTGCTGAACTGCTGAA

4422

4361 GAAATGACGGCTGTGCTGAACTGCTGAACTGCTGAACTGCTGAACTGCTGAA

4302

721 TATGTTGAACATTGCACTTGGTAACTGCGCTGAACTGCTGAACTGCTGAACTGCTGAA

720

4301 TATGTTGAACATTGCACTTGGTAACTGCGCTGAACTGCTGAACTGCTGAACTGCTGAA

780

781 TGAAGGTTATGCTCTGGTATGCTGAAATTATGGCTTCCACCGTCAAAAAA

840

4241 TGAGGTTATGCTCTGGGGAATCTGAGGATTATGGCTTCCACCTTCACCTCAAAA

4182

841 AACGGCGCTTGGCGCTTATTTCAACCTTATTCCACATACGACTACATAC

900

4181 AACGGCGCTTGGCGCTTATTTCAACCTTATTCCACATACGACTACATAC

4122

901 GTCCGTTTAACTCTTAACTGCTT 927

4121 GTCCGTTTAACTCTTAACTGCTT 4095

RESULT 5

AP002569/c

AP002569

AP002569

AP002569 BA00007

1 GI:13364704

DEFINITION

Escherichia coli O157:H7 DNA, complete genome, section 20/20.

VERSION

AP002569_1

KEYWORDS

Escherichia coli O157:H7

ORGANISM

Bacteria: Proteobacteria; Gammaproteobacteria; Enterobacteriales:

Enterobacteriaceae; Escherichia.

REFERENCE

1

Makino, K., Yokoyama, K., Kubota, Y., Yutsudo, C.H., Kimura, S.,

Kurokawa, K., Ishii, K., Hattori, M., Tatsuno, I., Abe, H., Iida, T.,

Yamamoto, K., Ohnishi, M., Hayashi, T., Yasunaga, T., Honda, T.,

Sasaki, C., and Shingawa, H.

TITLE

Complete nucleotide sequence of the prophage *Vtp2*-Sakai carrying the *Vtp2* genes of the enterohemorrhagic *Escherichia coli* O157:H7 derived from the Sakai outbreak.

JOURNAL

Genes Genet. Syst. 74 (5), 227-239 (1999)

MEDLINE

2019870

PUBMED

20734605

c013_0157:H7 ortholog: z5927"

/codon_start=1

/transl_table=11

/product="Icosapartyl dipeptidase"

/protein_id="AAH83830_1"

/db_xref="GI:2611643"

/translation="MIDYSAAFTLQGAMHAPDRCGDVILANGKIIAVASNIPS

DIVPDCWVDISGQIICPGFTDQHFLIGGEAEPTTRPEVALSRUITEAGTTSVG

LIGTDJSRHSPESSLAKTRAINBEGISAWMLTGATHVRSPRTGSVEKOVAIDRTVG

VKAISHSRSPADPVHLANNAEBSRGVLGGKGKVTFEHMDGSKALQVIDLLEN

LARVTSDDGNSQPFDDSGNLTHIGVACFTETLVQVLDYSESLSDAWRPLPS

SVAGFLNDIGKEBLUPGNDADLWMPERIEQTYARGKLMVUDKGKACVKGTETA"

Query Match 98.4%; Score 912.6; DB 1; Length 86898;

Matches 918; Conservative 0; Pred. No. 5.3e-260; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GTCATGTTAACGGGATCTGTTAATTGCAAGCATGCATGTCAGCAGTCAGGACTTGT 60

Db 85652 GTCATGTTAACGGGATCTGTTAATTGCAAGCATGCATGTCAGCAGTCAGGACTTGT 85593

Qy 61 ACTTCCCTGTTGATTAGTAGCGCAATTAGTGAGCAAAATGAGACCCGACATCT 120

Db 85592 ACTTCCCTGTTGATTAGTAGCGCAATTAGTGAGCAAAATGAGACCCGACATCT 85533

Qy 121 TATCGTGAAGAGCAGCTGGTAAACAGCACACGTGAGAAAGTATTGGAGGGAAAAG 180

Db 85532 TATCGTGAAGAGCAGCTGGTAAACAGCACACGTGAGAAAGTATTGGAGGGAAAAG 85473

Qy 181 CTATGAGTTTCGAGGAGCAGAGGGGAAATGATCATGAGTCATCTCTCTGAATATGA 240

Db 85472 CTATGAGTTTCGAGGAGCAGAGGGGAAATGATCATGAGTCATCTCTCTGAATATGA 85413

Qy 241 CATCAACCTGTTGATCATGATCAATCTGCCCGTAGAAGAGGTCTCTGTAGGGG 300

Db 85412 CATCAACCTGTTGATCATGATCAATCTGCCCGTAGAAGAGGTCTCTGTAGGGG 85353

Qy 301 TGAATCGGCGAGCAGGGAAATGTTGGGTATGTTCTGACTTGCGGTGAGCAAGGT 360

Db 85352 TGAATCGGCGAGCAGGGAAATGTTGGGTATGTTCTGACTTGCGGTGAGCAAGGT 85293

Qy 361 CGATAAATTCTGGCTCGAAATGGTCCAGTACATACCAAAACGGTTAACCG 420

Db 85292 CGATAAATTCTGGCTCGAAATGGTCCAGTACATACCAAAACGGTTAACCG 85233

Qy 421 GCGTGTACTTCAAGTCTGTCAGCAACTACTGTCGGTACGATCTGGTACTGT 480

Db 85232 GCGTGTACTTCAAGTCTGTCAGCAACTACTGTCGGTACGATCTGGTACTGT 85173

Qy 481 CAGGGAGAGCACTGTTGAGCTGAACTCTACAGTCAATGTTGGAACCTGGACATCA 540

Db 85172 CAGGGAGAGCACTGTTGAGCTGAACTCTACAGTCAATGTTGGAACCTGGACATCA 85113

Qy 541 CAACCGGTTGTTGATCTGCCCTGATGGCGGAACTGAGTCAGGTGGTCCCG 600

Db 85112 CAACCGGTTGTTGATCTGCCCTGAGTCAGGTGGTCCCG 85053

Qy 601 CACCCATGCTTCACTCTCTGAAACCCAGGAAATTCTAGTCCCCTGTAAGCTGTAA 660

Db 85052 CACCCATGCTTCACTCTCTGAAACCCAGGAAATTCTAGTCCCCTGTAAGCTGTAA 84993

Qy 661 GAAATGACCGCCCGTCAACCGCAAGCAGCTGTTGAGTCAGATGCCCG 720

Db 84992 GAAATGACCGCCCGTCAACCGCAAGCAGCTGTTGAGTCAGATGCCCG 84933

Qy 721 TATTCGTTAACATTGAACTTACCGGAAATCTACGCCGAAATCTACGCCGAACTTCGG 780

Db 84932 TATTCGTTAACATTGAACTTACCGGAAATCTACGCCGAAATCTACGCCGAACTTCGG 84873

Qy 781 TGAAGTTATGCTCTGCGGTGAGGATTAATGGTTACCGCTTACCAAAA 840

Db 84872 TGAAGTTATGCTCTGCGGTGAGTCAAGGAAATTCGGTTACCCGTCATAA 84813

Qy 841 AACGGCGCTTTAGGCCGTTTATTTCAACCTTTCAGTACGTAACTC 900

Db 84812 AACGGCGCTTTAGGCCGTTTATTTCACTTATTCAGTACGTAACTC 84753

Qy 901 GTCCGTGTAACCTCTTACTGCCTT 927

Db 84752 GTCCGTGTAACCTCTTACTGCCTT 84726

RESULT 7 ECODE ECODE E.Coli dye gene coding for Dye protein, complete cds. linear BCT 26-APR-1993

DEFINITION E.coli dye gene coding for Dye protein, complete cds.

VERSION M10044 GI:145817

KEYWORDS dye gene.

SOURCE Escherichia coli

ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.

REFERENCE J. Biol. Chem. 260 (7), 4236-4242 (1985)

AUTHORS Drury, L. S. and Buxton, R.S.

TITLE DNA sequence analysis of the dye gene of Escherichia coli reveals amino acid homology between the dye and OmpR proteins

JOURNAL J. Biol. Chem. 260 (7), 4236-4242 (1985)

MEDLINE 85157583

PUBMED 2984198

COMMENT Original source text: E.coli K12 DNA, clone PRB52.

Draft entry and sequence in computer readable form kindly provided by R.S.Buxton, 01-AUG-1985 [1].

Mutation of the dye gene results in sensitivity to dyes, envelope protein changes, loss of expression of alkaline phosphatase, and reduced transcription of sex factor F genes. The positions of the dye and OmpR genes on the K12 map are indicative of gene duplication and the Dye protein is 28% homologous to the OmpR protein. A potential transcription termination site is located at positions 606-940 and a ribosome binding site at 161-166. An URF was found on the complementary strand at positions >1542 to 948 (AA at 1541). location/Qualifiers

FEATURES source

gene /organism="Escherichia coli"
mol_type="genomic DNA"
/db_xref="taxon:562"
/gene="dye"
/gene="dye"
/gene="dye"

CDS
/gene="dye"
/codon_start=1
/transl_table=11
/protein_id="AAH83830_1"
/db_xref="GI:45818"
/translation="MOTPRILIVEDELVNTNLSIFEAEGYDVFEATDGARNHOIIS
KPNFRELTRARNLRSNGLLIREQANMLFLGRDNWDKINGLLEGADDYT
LPSEFRAMLFCEPNPKIOSRABLLKMTGKBLPHDRTVDTIRRKHFESTDT
PEIAIMHGFYRFGTCD"

ORIGIN PEIAMIHGFYRFGTCD

Query Match 98.4%; Score 912; DB 1; Length 1468;

Matches 923; Conservative 0; Pred. No. 3.7e-260; Mismatches 0; Indels 1; Gaps 1;

Qy 5 TGTAGCCGATCTGTTATTGAGCTGATGGCTGAGCGCACTGC 63

Db 1 TGTAGCCGATCTGTTATTGAGCTGATGGCTGAGCGCACTGC 84873

Qy 6 TCTGTTCAATTAGTGCAATTAGTGAGCACTGAGCAGCTTGTACT 60

Db 61 TCTGTTCAATTAGTGCAATTAGTGAGCACTGAGCAGCTTGTACT 120

Qy 124 CGTGAAGCGGAGTGTGAGCAACGCAACGTTGAGAATTTGGAGGAAAGCTA 183

| | | TITLE | Oiang, B.O., Wen, Y.M., Hou, Y.D. and Yu, J. |
|---|-----------------|---|---|
| | | JOURNAL | Genome sequence of <i>Shigella flexneri</i> 2a: insights into pathogenicity through comparison with genomes of <i>Escherichia coli</i> K12 and O157 |
| | | PUBLISHED | Nucleic Acids Res. 30 (20), 4432-4441 (2002) |
| Db | 121 | CGTGAAGAGGACTGTGGTACAGCAACAGTGAAAGTATTTCGAGGGAGGTA | 180 |
| Oy | 184 | TGATGTTTCGAGGCAGCATGGCAGATGGCGCGGAATGCACTGATGATCCTCTTGAAATGACAT | 243 |
| Db | 181 | TGATGTTTCGAGGCAGCATGGCAGATGGCGCGGAATGCACTGATGATCCTCTTGAAATGACAT | 240 |
| Oy | 244 | CAACCTGGTGTATGATCACTGCGGCTAAGAACGCTCATGATGATCCTCTTGAAATGACAT | 303 |
| Db | 241 | CAACCTGGTGTATGATCACTGCGGCTAAGAACGCTCATGATGATCCTCTTGAAATGACAT | 300 |
| Oy | 304 | ACTCGCCGAGCAGCGAATGTTGGTGTATGTTCTGACTGGCGTGAACAGAAGTGA | 363 |
| Db | 301 | TAATAATTCTGGCCTGAAATCGTGCAGATGACTACATACCAACCGTCAACCGCG | 423 |
| Oy | 364 | TAATAATTCTGGCCTGAAATCGTGCAGATGACTACATACCAACCGTCAACCGCG | 360 |
| Db | 361 | TAAATTCTGGCCTGAAATCGTGCAGATGACTACATACCAACCGTCAACCGCG | 420 |
| Oy | 424 | TGAACTGAGGATGTTGCAACGCAACTACTGTCGCCACATGAACTCGGGTACTGTCAG | 483 |
| Db | 421 | TGAACTGAGGATGTTGCAACGCAACTACTGTCGCCACATGAACTCGGGTACTGTCAG | 480 |
| Oy | 484 | CGAGAACGCTGAGCGTGAAGACTACAGTCAAGTCAATGTTGGGAACTGGACATCAACAG | 543 |
| Db | 481 | CGAGAACGCTGAGCGTGAAGACTACAGTCAATGTTGGGAACTGGACATCAACAG | 540 |
| Oy | 544 | CCGTTGTTGATGGGCCATGCGGAGCGTGAACCGCAGCGAGTCCCGC | 603 |
| Db | 541 | CCGTTGTTGATGGGCCATGCGGAGCGTGAACCGCAGCGAGTCCCGC | 600 |
| Oy | 604 | CAGCTTCACTCTGTGAAACCCAGGCAAATTCACTGTCGCCCTGAACTGTCAG | 663 |
| Db | 601 | CAGCTTCACTCTGTGAAACCCAGGCAAATTCACTGTCGCCCTGAACTGTCAG | 660 |
| Oy | 664 | AATGACCGCCGCGTGAAGCTGAACCGCAGCGAGTCCCGC | 723 |
| Db | 661 | AATGACCGCCGCGTGAAGCTGAACCGCAGCGAGTCCCGC | 720 |
| Oy | 724 | TGCTAACATTTGATCTACGGGAAATCGCCACCATCACGGGA | 783 |
| Db | 721 | TGCTAACATTTGATCTACGGGAAATCGCCACCATCACGGGA | 780 |
| Oy | 784 | AGCTTATGCTCTGCGATCTGGAAATATGCTTACACCGTAAAAAAC | 843 |
| Db | 781 | AGCTTATGCTCTGCGATCTGGAAATATGCTTACACCGTAAAAAAC | 840 |
| Oy | 844 | GCGCTTTTGGCGGTTTATTTAACCTTTCAGACATACACTATGTC | 903 |
| Db | 841 | GCGCTTTTGGCGGTTTATTTAACCTTTCAGACATACACTATGTC | 900 |
| Oy | 904 | CGTGTAACTCTTACTGGCTT | 927 |
| Db | 901 | CGTGTAACTCTTACTGGCTT | 924 |
| RESULT 8 | | | |
| AE01548/c | | | |
| DEFINITION | AE01548 AE01548 | 11861 bp DNA linear BCT 18-OCT-2002 | |
| ACCESSION | AE01548 | <i>Shigella flexneri</i> 2a str. 301 section 411 of 412 of the complete genome. | |
| VERSION | AE01548.1 | GI:24054959 | |
| KEYWORDS | | | |
| ORGANISM | | <i>Shigella flexneri</i> 2a str. 301 | |
| REFERENCE | | Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobactericeae; <i>Shigella</i> . | |
| AUTHORS | | Jin, Q., Shen, Y., Wang, J.H., Liu, H., Yang, J., Yang, F., Zhang, X.B., Zhang, J.Y., Yang, G.W., Zhao, A.L., Gao, Y.S., Zhu, J.P., Chen, S.X., Yao, Z.J., Wang, Y., Lu, W.C., Oiang, B.O., Wen, Y.M. and Hou, Y.D. | |
| FEATURES | | | |
| SOURCE | | | |
| LOCATION/QUALIFIERS | | | |
| 1. .11861 | | | |
| /organism="Shigella flexneri 2a str. 301" | | | |
| /mol_type="genomic DNA" | | | |
| /strain="301" | | | |
| /serotype="2a" | | | |
| /db_xref=taxon:198214 | | | |
| /complement=(185..1852) | | | |
| /gene=YJJK | | | |
| /locus_tag="SP4423" | | | |
| /gene=YJJK | | | |
| /note="Residues 1 to 555 of 555 are 99 pct identical to residues 1 to 555 of a 555 aa protein from <i>Escherichia coli</i> O157:H7 EDL933 ref: NP_291004.1" | | | |
| /codon_start=1 | | | |
| /transl_table=11 | | | |
| /product="putative ATP-binding component of ABC transport system" | | | |
| /protein_id="AAAN45837.1" | | | |
| /db_xref=GI:2405490 | | | |
| /translation="MADPVTMHRVKVVPKRLKLNLSFPGAKICGVGLGAG KSTLVKIMGDKOYEGEARFQDPIKIGVQFEPOLNPENTVRESEEVNVALK RLDEVALYVADPDDFKLAEQSLRLEIITODHNLYNLERADALLPWDKI ANLSCGERRVAVCFULLDPEMLDPHILDEVSVALERELPFLDPEVAWTHD RYFLDNVAGILEDGRGEFGFWEGNYSWMEQKDRQLAQASOBARRKISIEKELEW RQTKGRQSKKARLARFLNSTYKORENLNEFLIPPRPLGDVKULESNLRSYG DRLLDLSIPIKGAIVGIGNGAKSGKSLFMSQEDPSCTIGLTVEVLSRQF QFRDSMDNSKIVWEBSVGIDMKINGTEMPSRAVGRFFKGVDQGKRIGBLSGBR THILOQDKEVVEFEGNFTEYEYKKRTGADALEPKRIKYKRIAK" | | | |
| gene | | | |
| CDS | | | |
| 2033..4000 | | | |
| /gene="slc" | | | |
| /locus_tag="SP4424" | | | |
| /gene="slc" | | | |
| /locus_tag="SP4424" | | | |
| 2033..4000 | | | |
| /gene="slc" | | | |
| /locus_tag="SP4424" | | | |
| residues 1 to 654 of a 654 aa protein from <i>Escherichia coli</i> O157:H7 EDL933 ref: NP_291005.1" | | | |
| /codon_start=1 | | | |
| /transl_table=11 | | | |
| /product="soluble lytic murein transglycosylase" | | | |
| /protein_id="AAAN45838.1" | | | |
| /db_xref=GI:24054911 | | | |
| /translation="MMLHEDALVEKAQVWTLAAGULTVSSVARDSLDEQR SYAQIKQDNDRNQDMVDEMMPGKDYPPVYLPYRQTDLMQPAVVTNFRA PTKLPPARTPSQPNACDPSVNLAREDDWRDQLLSPSKERPTTEACQNYYAKWNGQSQEWAQ PAKEWLWTGSOPNACDPSVNLAREDDWRDQLLSPSKERPTTEACQNYYAKWNGQSQEWAQ ADYOTIASATISLANNPNTLIFARTGDFTRONAAVAFASVQDANRJLMP LAQAOQNLNEQDQBLERDWTWRLMNDVDEQBQKDDA1MRSOSTSLRRVMAV TGDREGINTMLARMEKDRDYYWQDYLRLGRGEBAKEIHLIOJQREPMV AVORIGEEVYKLKIDAPONVDSLTIQGPENARVELMYMNDNTARSENWLNVSK TEQDQLARYVNPNQWPSLDSQTATGKWLHLSRPFPLANDPFLRPTYSGKEFQSY MAJQOESAWNPKVQWPSLDSQTATGKWLHLSRPFPLANDPFLRPTYSGKEFQSY GTSIYQVQFGNRFISSAYNAGPGRTRWLGNSAGTIDAVAFESPPSETGY VNRVLAYDAYRYRFMWDKPLMSATEWGRYY" | | | |

gene 4090..4416
/CDS
/locus tag="SF4425"
/locus tag="trpR"
/locus tag="SF4425"
/note="Residues 1 to 108 of a 108 aa protein from Escherichia coli O157:H7 EDL933 ref: NP_291061.1"
/codon start=1
/transl_table=11
/product="regulator for trp operon and aroH; trp opererepressor"
/protein id="AN45839_1"
/db xref="GI:2405462"
/translation="MAQSPSPSAMAEQRHOEWLRFVFLKNAVONDHLPLNLMLTPIERALGTRVRIIVEELLERGENSQRELNKELGAGATITRGSNSLKAPELQWLEEVLKSD"
/gene=YjjX
/CDS
/locus tag="SF4426"
/pseudo
/complement(4513..5030)
/locus tag="SF4426"
/note="similar to Escherichia coli K12 orf, conserved hypothetical protein gi: 1790855 (174 aa). BLAST with identity of 95% over 172 aa. This CDS contains frameshift. The sequence has been checked and is believed to be correct."
/pseudo
/codon start=1
/transl_table=11
5073..5720
/gene="gpmB"
/CDS
/locus tag="SF4427"
/gene="gpmB"
/locus tag="SF4427"
/note="Residues 1 to 215 of 215 are 100 pct identical to residues 1 to 215 of a 215 aa protein from Escherichia coli O157:H7 ref: NP_313380.1"
/transl_table=11
/product="phosphoglyceromutase 2"
/protein id="AA45840_1"
/db xref="GI:24054963"
/translation="MIQVYLFRGETOWAERRILOGOSDPSLTAKGEQAMQATRAKEIGGTHIISSDGLRTTAIAQCGCDIFDSRLEJMGVLEKRHIDSLIPEEENIIGLPAWAKERRKLRNCSIRVVDYQESLWASGVNTETAGDISHDAPDLELQR"
gene complement(5717..6586)
/locus tag="SF4428"
/complement(5717..6586)
/gene="rob"
/locus tag="SF4428"
/note="Residues 1 to 289 of 289 are 100 pct identical to residues 1 to 289 of a 289 aa protein from Escherichia coli O157:H7 EDL933 ref: NP_291009.1"
/transl_table=11
/product="right origin-binding protein"
/protein id="AN45841_1"
/db xref="GI:24054964"
/translation="MDQGIVRLDILWLEGHDQLPLSLDNVAAKAGYSKWHQLQRNFKDVTGAIAYTRABPLSKAVARLURPFLDIALQDFRSQDKDQEYFTTALAQOQADGVLTFGPVMYRREPMSAFGIPRPLRGEFTPEHEKVEITDPLIGVQSYSSLEOTSDFREMRVFFHDFLQNPHTPPVYGLNETRPSQDKDQEYFTTALAQOQADGVLTFGPVMQSGEYVMTFEGTGQVFELTFTVYCTCMPMLNLLRKGDIERVYPABDAKAGDRP6777..7270
/gene="creA"
/CDS
/locus tag="SF4429"
/locus tag="SF4429"
/note="Residues 1 to 157 of 157 are 100 pct identical to residues 1 to 157 of a 157 aa protein from Escherichia coli O157:H7 EDL933 ref: NP_291010.1"
/codon start=1
/transl_table=11
/product="orf, conserved hypothetical protein"
/protein id="AN45842_1"
/db xref="GI:24054965"
/translation="MKYKHILSLSLIMGPLAHAEIGSVDTVEKMGPDHKIVWEFDDPDVKNVYCYSVSKTGGIKSGIGLABDTSALISQQVGRIELSDRKGKACCEVVFKRTSTVLFKS1QVREVDKRNVALAYLSDKVVEGSPKNAISAVPMWQ"
/gene="creB"
/CDS
/locus tag="SF4430"
/gene="creB"
/locus tag="SF4430"
/note="Residues 1 to 229 of 229 are 98 pct identical to residues 1 to 229 of a 229 aa protein from Escherichia coli K12 ref: NP_418185.1"
/transl_table=11
/product="catabolic regulation response regulator"
/protein id="AN45843_1"
/db xref="GI:24054965"
/translation="MQEITLVVEDOGIAITVMLQDEGFATEVFGELPVLDKARQQFDVMLDVLDLUDPSGFCRLQALHPALPVLTARSEEDVDRDLCIGADYVAKOFSPRPVCVRVTRURVVKFESTPSPVIRIGHELNEBPAQOSWDFDPLTURYFHGMGTSLRKG"
/locus tag="SF4431"
/gene="creC"
/CDS
/locus tag="SF4431"
/gene="creC"
Query Match
Best Local Similarity 98.2%; Score 910.2; DB 1; Length 11861;
Matches 923; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 1 GTCACTGTAGCGCGATAATGTTAATGTCAGCATGCATTCAGCGAGTCGGACTTTCT 60
Db 11681 GTCATGTGTCGCGATCACTGTTATTGAGCATGTCATGGCAGGTCTGGACTTTCT 11622
QY 61 ACTCCCTTTGATTGTTGGCAATTAGTAGTCAACATGCAGACCCGCACATCT 120
Db 11621 ACTTCCTGTTGATTGTTGGCAATTAGTAGTCAACATGCAGACCCGCACATCT 11562
QY 121 TATCGTTGAACAGCTTGTGTACAGCAACACGTGAAAGATTGGAACCGGAAG 180
Db 11561 TATCGTTGAACAGCTTGTGTACAGCAACACGTGAAAGATTGGAACCGGAAG 11502
QY 181 CTATGATGTTTGTGAGCCAGCATGGCGGAGATGTCAGCATCCCTCTGAATATGA 240
Db 11501 CTATGATGTTTGTGAGCCAGCATGGCGGAGATGTCAGCATCCCTCTGAATATGA 11562
QY 241 CATCACCTGTTGATCATGATCACTGCGGAAAGACGGCTTCTGTAGCG 300
Db 11441 CATCACCTGTTGATCATGATCACTGCGGAAAGACGGCTTCTGTAGCG 11382
QY 301 TGAACCTGCCGAGCAGCGATGTGCTTGTGTTCTGACTGCGCTTCAAGAAGT 360
Db 11381 TGAACCTGCCGAGCAGCGATGTGCTTGTGTTCTGACTGCGCTTCAAGAAGT 11442
QY 361 CGATTAATCTCGCCCTCGAATCTCGCTGCGATGACTTCAACCGTCAACC 420
Db 11321 CGATAAATCTCGCCCTCGAATCTCGCTGCGATGACTTCAACCGTCAACC 11262
QY 421 GCGTGAAGTGAAGATTCGTCAGCACTACTGCCCCAACATGATCTGGTACTGT 480

Db 11261 GCGTGAACGTGACGATTCTGCAGCAACCTGCTGCCGTTACATGAATCTGGTACTGT 11202 CDS
 Qy 481 CAGCGAAGAACGTCGAGCCTGAAACSTCAAGTCAATGGTGGAAACTGACATCAA 540
 /gene="fimb"
 Db 11201 CAGCGAACAAAGCTGAGCTGAGCTGAAAGCTCAAGTCAGTTCAATGGTGGAAACTGACATCAA 11142
 541 CAGCGGTCTGGTATCGCCCTGATGGGGAGCTACAGTCAATGGTGGAAACTGACATCAA 540
 /locus_tag="S4466"
 /functions="regulator; Surface structures"
 /note="residues 1 to 198 of 198 are 100.00 pct identical to residues 1 to 198 or 198 from Escherichia coli K-12 : B4313; involved in phase variation"
 /codon_start=1
 /transl_table=11
 /product="recombinase; regulator for fimb"
 /protein_id="AAP1945.1"
 /db_xref=GI:3004366
 /translation="MSKRRVITKGKEVQAMQAVCNGATGARDYCHILLAVRHGMRISE LUDHYDULNEGRNIRLKLNGGSTVHLRFDEBEAVRWTOSERANWKGADRTDAI FISRGSRSLRQAIVTRIAGIAGTQTHMLRHAGCYELAERGADTRLIQDYL GHNRNIKHTVRITASNAFAGIWERNLINNEBLKREEV"
 complement(1462..1947)
 /gene="fimb"
 /locus_tag="S4467"
 /note="residues 1 to 160 of 199 are 97.50 pct identical to residues 1 to 160 of 200 from Escherichia coli K-12 : B4312; involved in phase variation"
 /codon_start=1
 /transl_table=11
 /product="recombinase; regulator for fimb"
 /protein_id="AAP16.1"
 /db_xref=GI:3004367
 /translation="MRNKANKKRNFLTHESEISLKAANTGPHTATNYQULGFIH GFRASICRISIDTDLACACIIRLKPKRSTPRLNVEOALKWMSIRTSPHA ESEWEFLSRKQNPLRQFPHIISTSGGNAGLSLEIHPHMLRYSGCFALNMGIDRL I"
 gene
 3569..4117
 /gene=YJ12A
 /locus_tag="S4468"
 3569..4117
 /gene=YJ12A
 /locus_tag="S4468"
 /note="residues 1 to 182 of 182 are 100.00 pct identical to residues 60 to 241 of 241 from Escherichia coli K-12 : B4311"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="AAP1417.1"
 /db_xref=GI:3004368
 /translation="MESENNTNTIHONKKENNAALNDVOVENEYAIKDDDMTPVRGMLT HFSNSGTRYKVKLSDATKDLDLNGCIRYKWDKAYRQDLSGMSDRDNHRDGYMT YHINSUFTFWQTTLXSKQDYRYANKKRATENAFLQHMTDPTPYBYDYLQRQ GYVNGRDNLTEENSYRIGVSPKL"
 4029..523
 /gene=YJ12B
 /locus_tag="S4469"
 4029..523
 /gene=YJ12B
 /locus_tag="S4469"
 /note="residues 1 to 404 of 404 are 98.76 pct identical to residues 1 to 404 of 404 from Escherichia coli K-12 : B4310"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="AAP1948.1"
 /db_xref=GI:3004369
 /translation="MWTIILTRVRFPAITIYKTVIALVCHLNCROBTWVNMKTIAL AMMSSFAANASVLPETPQKSCGAIIDWYIGSCTAWXKLUOTWAKDQKWTIA LAFFPGGPQRQTSARTIDGALYVFGEGKNSBGLIQVNFDVHKAPTKMSWKLMASHA PMGMAGCHVTFVINGKAVYTGQNQNFNGYFEDLINEAGKUSTAIDKINAHYFDKIAED
 FEATURES
 source
 1..22594
 /organism="Shigella flexneri 2a str. 2457T"
 /mol_type="genomic DNA"
 /strain="2457T"
 /serotype="2a"
 /ab_xref="Taxon:198215"
 complement(271..867)

YFENKELLSRDPSTQWSVAGEPWPYCTAGAAVWNKDKTMILNGRAKPGRLTAVPE
LDFDTNLNWKWKGUDPVSSPDGVWGGPAGISNDLIFGGAFPRKGSSENYQNGKNAHE
GLKSYSTIDHILWINGKWDKXSELQSQRAGVSLPWNNSLLIGGETAGSKAVTDSSL
ISVKDVKVTON" 5530. . 6288

gene /locus tag="X1hs" /codon start=1
/product="hypothetical protein"
/db_xref="GI:30043703" /transl_table=11
/translation="MGGVHPALATNHOYGTGQALHTARKLKPFTDNAGILIVPC
WPHSGERDLMSDYASHQHNNMWEAFRDQLKQHSQDNITPAWFCGDTWVNPW
PPHSEVAKIGMUNVNLANTIVDFOQOQERGLQHNPDDPLSTGYGSAYRSPE
WTALKSSHSTAARRGIISDRFVEALQFWRR" 6742. . 7509

repeat_region /notes="residues 1 to 252 of 252 are 99.60 pct identical to residues 75 to 326 of 326 from Escherichia coli K-12."
/codon start=1
/product="hypothetical protein"
/db_xref="GI:30043700" /transl_table=11
/translation="MGGVHPALATNHOYGTGQALHTARKLKPFTDNAGILIVPC
WPHSGERDLMSDYASHQHNNMWEAFRDQLKQHSQDNITPAWFCGDTWVNPW
PPHSEVAKIGMUNVNLANTIVDFOQOQERGLQHNPDDPLSTGYGSAYRSPE
WTALKSSHSTAARRGIISDRFVEALQFWRR" 6742. . 7260

gene /insertion_seq="ISX1" complement(6757. . 7260)
/locus tag="S4471" complement(6757. . 7260)

CDS /locus tag="S4471" /function="IS, phage, Tn; Transposon-related functions"
/notes="residues 1 to 167 of 167 are 89.82 pct identical to residues 1 to 167 of 167 from GenPept : >gb|AF414742.1|AF414323_13 (AF414323) Tnph [Shigella flexneri]" /codon_start=1
/transl_table=11
/product="IS1 orfB"
/protein_id="AP414742.1"
/db_xref="GI:30043701" /translation="MSSQCTHGRWPOHGFSLKKLRPQSVTSRQIOPGSVDIVVACMD
WPIYESCIRKGLKIVVISRKTERINLNHQHLARLVRSLSTSRSKSVLHDKVIGH
LNUKHYQ" /complement(7179. . 7454)
/locus_tag="S4472" complement(7179. . 7454)

gene /locus tag="S4472" /function="IS, phage, Tn; Transposon-related functions"
/notes="residues 1 to 91 of 91 are 98.90 pct identical to residues 1 to 91 of 91 from GenPept : >gb|AF414850.1|AF348706_239 (AF348706) iso-IS1 orfA" /codon_start=1
/transl_table=11
/product="IS1 orfA"
/protein_id="AP414850.1"
/db_xref="GI:30043702" /translation="MASICRCPSATEGVVNRKGKSTAGHLRLCSHCRKTWQQT
YTASORGHRKIKMAMNGVGCRASARINGVGLNTVLRHLKNSGRSR" 7549. . 8244

gene /locus tag="S4473" /codon_start=1
/product="putative regulatory protein, deoR family"
/db_xref="GI:30043703" /translation="MASICRCPSATEGVVNRKGKSTAGHLRLCSHCRKTWQQT
YTASORGHRKIKMAMNGVGCRASARINGVGLNTVLRHLKNSGRSR" 7549. . 8244

CDS /locus tag="S4473" /function="regulatory"
/notes="residues 1 to 231 of 231 are 81.46 pct identical to residues 30 to 261 of 261 from GenPept : >gb|AF4142652.1| (Salmonella typhimurium LT2)" putative regulatory protein, deoR family
/codon_start=1 /transl_table=11
/product="putative deoR family transcriptional regulator"
/protein_id="AP4142652.1" 541 CAGCCCTTCGTTATGGCCCTGAGCGAGCTACAACCTCATGGCTTCACTGAGCTGAGCTGGTCCG 600
Db 224158 CAGCGTTCCTGGCCCTGAGCGAGCTACAACCTCATGGCTTCACTGAGCTGAGCTGGTCCG 600
Qy 601 CGCCATGCTCACTCTGAGAACCCAGCAAAATTCAATGCCGCTGACTGACATCA 540
Db 224218 CGCGAGAACGCTCTAGCTGAGCTACAACCTCATGGCTTCACTGAGCTGGTCCG 660
Qy 661 GAAATGACGGCGTGAAGAACCGAACCGAACCGAACGACTGAGCTGAGATCGGCC 720
Db 2244038 GAAATGACGGCGTGAAGAACCGAACCGAACCGAACGACTGAGCTGAGATCGGCC 223979
Qy 721 TATCGTAAACATTCTGAACTACGCCGATAGCCGAGATCGCCACCATCAGCG 780

| | | | |
|------------|--|---|----------|
| Db | 223978 | TATTCGTAACATTGAACTCAGCCGATAGCCGGATATCGCCGATATCTGCACCAATTAGG | 223919 |
| Qy | 781 | TGAAGTGTATCGCTTCGGTGTATCTGGAGATAATCGGTATACCGCGTAAANAA | 840 |
| Db | 223918 | TGAGGTATCGCTTCGGTGTATCTGGAGATAATCGGTATACCGCGTAAANAA | 223860 |
| Qy | 841 | AACGGGCTTTAGCGGTTTATTCACACTTATCCAGATAGTACTCATC | 900 |
| Db | 223859 | AACGGGCTTTAGCGGTTTATTCACACTTATCCAGATAGTACTCATC | 223800 |
| Qy | 901 | GRCGGTGTAACTCTTACTGCCTT | 927 |
| Db | 223799 | GRCCGTGTTACTCTTACTGCCTT | 223773 |
| RESULT 10 | | | |
| LOCUS | AL627284/c | AL627284 | 88037 bp |
| DEFINITION | Salmonella enterica serovar Typhi (Salmonella typhi) strain Ctl8, | | |
| ACCESSION | complete chromosome; segment 20/20. | | |
| KEYWORDS | AL627284.1 GI:16505618 | | |
| SOURCE | Salmonella enterica subsp. enterica serovar Typhi | | |
| ORGANISM | Bacterium; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonellae. | | |
| REFERENCE | I (bases 1 to 88037) | | |
| AUTHORS | Parkhill,J., Dougan,G., James,K.D., Thomson,N.R., Pickard,D., Wain,J., Churcher,C., Mungall,K.L., Bentley,S.D., Holden,M.T.G., Sebbah,M., Baker,S., Basham,D., Brooks,K., Chillingworth,T., Connerton,P., Cronin,A., Davis,R.M., Davies,R.M., Dowd,L., White,N., Farrar,J., Feltwell,T., Hamlin,N., Haque,A., Hien,T.T., Holroyd,S., Jagels,K., Krogh,A., Larsen,T.S., Leather,S., Moulie,S., O'Gaora,P., Parry,C., Quail,M., Rutherford,K., Simmonds,M., Skelton,J., Stevens,K., Whitehead,S. and Barrell,B.G. | | |
| TITLE | Complete genome sequence of a multiple drug resistant <i>Salmonella enterica</i> serovar Typhi Ctl8 | | |
| JOURNAL | Nature 413 (6858), 848-852 (2001) | | |
| PUBMED | 11677608 | | |
| REFERENCE | 2 (base 1 to 88037) | | |
| AUTHORS | Parkhill,J. | | |
| TITLE | Direct Submission | | |
| JOURNAL | Submitted (25-OCT-2001) Submitted on behalf of the <i>Salmonella</i> sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK | | |
| E-mail: | Parkhill@sanger.ac.uk | | |
| Notes: | Details of <i>S. typhi</i> sequencing at the Sanger Centre are available on the world wide Web. (URL, http://www.sanger.ac.uk/Projects/S_typhi/). | | |
| FEATURES | | | |
| Source | | | |
| gene | /organism="Salmonella enterica subsp. enterica serovar Typhi" /mol_type="genomic DNA" /strain="Ctl8" /db_xref="taxon:90370" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="CA021411" /db_xref="GI:16505619" /ab_xref="EMBL:Q8ZQY5" | | |
| CDS | | | |
| gene | /genes="STY4858" /genes="STY4859" /note="Similar to <i>Streptomyces coelicolor</i> hypothetical protein SCMH1.19 TR:054178 (EMBL:AL021411) (182 aa) Fasta scores: E(): 1.4e-10, 29.6% id in 169 aa" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="CA03347.1" /db_xref="GI:16505619" /ab_xref="EMBL:Q8ZQY5" | | |
| CDS | | | |
| gene | /genes="STY4859" /note="N-terminal signal sequence" /codon_start=1 /transl_table=11 /product="putative exported protein" /protein_id="CA03349.1" /db_xref="GI:16505621" /transl_table="MSTIKPFLAVIMLILIRPGATEDGAIWVKTISAYDYLDTARLHVADIGLVUGBFHDARAQNVLKPPPTVLFVGNPKGGTPIMLAHPPELDLPRVILSQADGRTLYPSRAETIQLQRGDAIQAKKLEQLOVEKSLH" complement(3051 .. 3716) /codon_start=1 /complement(3051 .. 3716) /gene="STY4861" /note="Similar to <i>Bacillus subtilis</i> hypothetical protein YCZH YCZH_YCZH_BAGSU (O31482) (185 aa), Fasta scores: E(): 7.4e-16, 31.4% id in 159 aa" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="CA03350.1" /db_xref="GI:16505622" /ab_xref="GI:0820Y2" /db_xref="EMBL:Q8ZQY2" /translation="MKDTVFFASVSGGSNASSSSRRRAILHEINGINAHQAGHEWWTRGFOLYYPLPHNPFRPFREQEBCAYRHRSANVFDPAVWTLHDHQYETLIVGYSVATAMLSAASGLCIVCIVGYSRROYHLAPLCALVARYEAFDRAMQTLERPRVQCKMDARGFCDDASATFADALSHQVMDVSAFIRDITCANTSPD" | | |

| | misc_feature | Query Match | 77.2% | Score | 715.6; | DB | 1; | Length | 88037; | | |
|---------------|--|-----------------------|-------|-------|----------------|------------|----|--------|--------|------|----|
| | gene | Best Local Similarity | 92.2% | Pred. | No. 2..9e-201; | Mismatches | 0; | Indels | 0; | Gaps | 0; |
| 4066 . . 4839 | /note="uxur" | | | | | | | | | | |
| 4066 . . 4839 | /note="synonym: STY4861" | | | | | | | | | | |
| 4066 . . 4839 | /note="PS00120 Lipases" | | | | | | | | | | |
| 4066 . . 4839 | /note="uxur" | | | | | | | | | | |
| 4066 . . 4839 | /note="similar to Escherichia coli uxu operon transcriptional regulator uxu_ecoli (P39161) (257 aa) fasta scores: E(): 0, 89.9% id in 257 aa, and to Erwinia chrysanthemi exu regulon transcriptional regulator exur SW:EKUR_ECOLI (Q9X9G0) (259 aa) fasta scores: E(): 0, 64.9% id in 256 aa. Consid residues 34 to 55, score 975(+2.51 SD). Fasta hit to PDHR_ECOLI (254 aa), 32% identity in 221 aa overlap | | | | | | | | | | |
| 4066 . . 4839 | Fasta hit to EKUR_ECOLI (258 aa), 48% identity in 255 aa overlap | | | | | | | | | | |
| 4066 . . 4839 | Orthologue of E. coli uxur (UXUR_ECOLI); Fasta hit to UXUR_ECOLI (257 aa), 90% identity in 257 aa overlap | | | | | | | | | | |
| 4066 . . 4839 | /cood_start=1 | | | | | | | | | | |
| 4066 . . 4839 | /transl_table=11 | | | | | | | | | | |
| 4066 . . 4839 | /product="uxu operon transcriptional regulator" | | | | | | | | | | |
| 4066 . . 4839 | /protein_id="CA03351.1" | | | | | | | | | | |
| 4066 . . 4839 | /db_xref="GI:6505623" | | | | | | | | | | |
| 4066 . . 4839 | /db_xref="TREMBL:Q8Z0Y1" | | | | | | | | | | |
| 4066 . . 4839 | /translation="MKSNTSQQRYYQEVAAMIRDLIVDVPYRPERLPPPERETARIN VTRVVRALIMIPLKIGLKVWILSDANNNEGADYNHNCNDAGPFELQLQR QLEESNTIAFAALQATREDIKMQQLAQELRSASSRGGSQSDDMOPHLATEATH NSMVEELFRQSQRRENENMWLQLQSHGDLYKEREWGDHKRQILLAALIKKDAAKL AMMQLHENKTQRLIEFSPNVDIYFDGLYLFESWPDNVDA" | | | | | | | | | | |
| 4066 . . 4839 | /gene="uxur" | | | | | | | | | | |
| 4066 . . 4839 | /note="Fram match to entry PF00392 gntR, Bacterial regulatory proteins, gntR family, score 107.20, E-value 1.4e-31" | | | | | | | | | | |
| 4066 . . 4839 | /gene="uxur" | | | | | | | | | | |
| 4066 . . 4839 | /note="PS00043 Bacterial regulatory proteins, gntR family signature" | | | | | | | | | | |
| 4066 . . 4839 | /gene="trps2" | | | | | | | | | | |
| 4066 . . 4839 | /note="Synonym: STY4863" | | | | | | | | | | |
| 4066 . . 4839 | /complement(4840. . 5853) | | | | | | | | | | |
| 4066 . . 4839 | /gene="trps2" | | | | | | | | | | |
| 4066 . . 4839 | /EC_number="6.1.1.2" | | | | | | | | | | |
| 4066 . . 4839 | /note="Similar to Clostridium longisporum tryptophanyl-tRNA synthetase Trps or Tiba SW:SYW_CLOLO (O46127) (341 aa) Fasta scores: E(): 0, 54.8% id in 330 aa, and to Thermotoga maritima tryptophanyl-tRNA synthetase Trps SW:SYW_THERA (Q9WYW2) (338 aa) Fasta scores: E(): 0, 36.5% id in 329 aa. Note that the E.coli trps ortholog is STY4315" | | | | | | | | | | |
| 4066 . . 4839 | /codon_start=1 | | | | | | | | | | |
| 4066 . . 4839 | /product="probable tryptophanyl-tRNA synthetase" | | | | | | | | | | |
| 4066 . . 4839 | /protein_id="CA03352.1" | | | | | | | | | | |
| 4066 . . 4839 | /db_xref="GOA:Q9B2Y0" | | | | | | | | | | |
| 4066 . . 4839 | /db_xref="TREMBL:Q8Z0Y0" | | | | | | | | | | |
| 4066 . . 4839 | /transl_table="MNUPLITLGDRPTGSLHLGHVYGSLRQRVQLQDHQHQVYLAD LOGDITLNGSNPKQKRTNIQVLAQDIAVLAQGIDPAFTTICLQSLAPALAAELTVLMMVNT VTRVKNPYVNEIQLRQGFTSPVQFCMAYTISQADITFARKAEMVPGDQDOLPMEQ AVSAMYTDPNHLKIDPCKLEGNVVUTWLDRAFPKAKVAKMANKQOGGIGDRCKN ELETIQELIAPIRERRATFADKSMMLMELUKKGSSERAHEVTOKTQLOBVKRGLGQPTL FOV" | | | | | | | | | | |
| 4066 . . 4839 | /complement(5170. . 5835) | | | | | | | | | | |
| 4066 . . 4839 | /note="PFam match to entry PE00579 tRNA-synt 1b_tRNA | | | | | | | | | | |
| 4066 . . 4839 | /note="trps2" | | | | | | | | | | |
| 4066 . . 4839 | /organism="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /source="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /version="29140243" | | | | | | | | | | |
| 4066 . . 4839 | /accession="AE016849" | | | | | | | | | | |
| 4066 . . 4839 | /locus="AE016849/c" | | | | | | | | | | |
| 4066 . . 4839 | /definition="Salmonella enterica subsp. enterica serovar Typhi Ty2, section 16 of 16 of the complete genome." | | | | | | | | | | |
| 4066 . . 4839 | /version="AE016849_1" | | | | | | | | | | |
| 4066 . . 4839 | /keywords="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /organism="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /source="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /version="29140243" | | | | | | | | | | |
| 4066 . . 4839 | /accession="AE016849" | | | | | | | | | | |
| 4066 . . 4839 | /locus="AE016849/c" | | | | | | | | | | |
| 4066 . . 4839 | /definition="Salmonella enterica subsp. enterica serovar Typhi Ty2, section 16 of 16 of the complete genome." | | | | | | | | | | |
| 4066 . . 4839 | /version="AE016849_1" | | | | | | | | | | |
| 4066 . . 4839 | /keywords="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /organism="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /source="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /version="29140243" | | | | | | | | | | |
| 4066 . . 4839 | /accession="AE016849" | | | | | | | | | | |
| 4066 . . 4839 | /locus="AE016849/c" | | | | | | | | | | |
| 4066 . . 4839 | /definition="Salmonella enterica subsp. enterica serovar Typhi Ty2, section 16 of 16 of the complete genome." | | | | | | | | | | |
| 4066 . . 4839 | /version="AE016849_1" | | | | | | | | | | |
| 4066 . . 4839 | /keywords="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /organism="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /source="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /version="29140243" | | | | | | | | | | |
| 4066 . . 4839 | /accession="AE016849" | | | | | | | | | | |
| 4066 . . 4839 | /locus="AE016849/c" | | | | | | | | | | |
| 4066 . . 4839 | /definition="Salmonella enterica subsp. enterica serovar Typhi Ty2, section 16 of 16 of the complete genome." | | | | | | | | | | |
| 4066 . . 4839 | /version="AE016849_1" | | | | | | | | | | |
| 4066 . . 4839 | /keywords="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /organism="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /source="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /version="29140243" | | | | | | | | | | |
| 4066 . . 4839 | /accession="AE016849" | | | | | | | | | | |
| 4066 . . 4839 | /locus="AE016849/c" | | | | | | | | | | |
| 4066 . . 4839 | /definition="Salmonella enterica subsp. enterica serovar Typhi Ty2, section 16 of 16 of the complete genome." | | | | | | | | | | |
| 4066 . . 4839 | /version="AE016849_1" | | | | | | | | | | |
| 4066 . . 4839 | /keywords="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /organism="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /source="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /version="29140243" | | | | | | | | | | |
| 4066 . . 4839 | /accession="AE016849" | | | | | | | | | | |
| 4066 . . 4839 | /locus="AE016849/c" | | | | | | | | | | |
| 4066 . . 4839 | /definition="Salmonella enterica subsp. enterica serovar Typhi Ty2, section 16 of 16 of the complete genome." | | | | | | | | | | |
| 4066 . . 4839 | /version="AE016849_1" | | | | | | | | | | |
| 4066 . . 4839 | /keywords="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /organism="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /source="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /version="29140243" | | | | | | | | | | |
| 4066 . . 4839 | /accession="AE016849" | | | | | | | | | | |
| 4066 . . 4839 | /locus="AE016849/c" | | | | | | | | | | |
| 4066 . . 4839 | /definition="Salmonella enterica subsp. enterica serovar Typhi Ty2, section 16 of 16 of the complete genome." | | | | | | | | | | |
| 4066 . . 4839 | /version="AE016849_1" | | | | | | | | | | |
| 4066 . . 4839 | /keywords="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /organism="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /source="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /version="29140243" | | | | | | | | | | |
| 4066 . . 4839 | /accession="AE016849" | | | | | | | | | | |
| 4066 . . 4839 | /locus="AE016849/c" | | | | | | | | | | |
| 4066 . . 4839 | /definition="Salmonella enterica subsp. enterica serovar Typhi Ty2, section 16 of 16 of the complete genome." | | | | | | | | | | |
| 4066 . . 4839 | /version="AE016849_1" | | | | | | | | | | |
| 4066 . . 4839 | /keywords="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /organism="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /source="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /version="29140243" | | | | | | | | | | |
| 4066 . . 4839 | /accession="AE016849" | | | | | | | | | | |
| 4066 . . 4839 | /locus="AE016849/c" | | | | | | | | | | |
| 4066 . . 4839 | /definition="Salmonella enterica subsp. enterica serovar Typhi Ty2, section 16 of 16 of the complete genome." | | | | | | | | | | |
| 4066 . . 4839 | /version="AE016849_1" | | | | | | | | | | |
| 4066 . . 4839 | /keywords="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /organism="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /source="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /version="29140243" | | | | | | | | | | |
| 4066 . . 4839 | /accession="AE016849" | | | | | | | | | | |
| 4066 . . 4839 | /locus="AE016849/c" | | | | | | | | | | |
| 4066 . . 4839 | /definition="Salmonella enterica subsp. enterica serovar Typhi Ty2, section 16 of 16 of the complete genome." | | | | | | | | | | |
| 4066 . . 4839 | /version="AE016849_1" | | | | | | | | | | |
| 4066 . . 4839 | /keywords="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /organism="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /source="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /version="29140243" | | | | | | | | | | |
| 4066 . . 4839 | /accession="AE016849" | | | | | | | | | | |
| 4066 . . 4839 | /locus="AE016849/c" | | | | | | | | | | |
| 4066 . . 4839 | /definition="Salmonella enterica subsp. enterica serovar Typhi Ty2, section 16 of 16 of the complete genome." | | | | | | | | | | |
| 4066 . . 4839 | /version="AE016849_1" | | | | | | | | | | |
| 4066 . . 4839 | /keywords="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /organism="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /source="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /version="29140243" | | | | | | | | | | |
| 4066 . . 4839 | /accession="AE016849" | | | | | | | | | | |
| 4066 . . 4839 | /locus="AE016849/c" | | | | | | | | | | |
| 4066 . . 4839 | /definition="Salmonella enterica subsp. enterica serovar Typhi Ty2, section 16 of 16 of the complete genome." | | | | | | | | | | |
| 4066 . . 4839 | /version="AE016849_1" | | | | | | | | | | |
| 4066 . . 4839 | /keywords="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /organism="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /source="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /version="29140243" | | | | | | | | | | |
| 4066 . . 4839 | /accession="AE016849" | | | | | | | | | | |
| 4066 . . 4839 | /locus="AE016849/c" | | | | | | | | | | |
| 4066 . . 4839 | /definition="Salmonella enterica subsp. enterica serovar Typhi Ty2, section 16 of 16 of the complete genome." | | | | | | | | | | |
| 4066 . . 4839 | /version="AE016849_1" | | | | | | | | | | |
| 4066 . . 4839 | /keywords="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /organism="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /source="Salmonella enterica subsp. enterica serovar Typhi Ty2 | | | | | | | | | | |
| 4066 . . 4839 | /version="29140243"</ | | | | | | | | | | |

| | | |
|---|---|-------------------------------------|
| 421 | GCCTGAACTGAGATCTGTGAGGAAACCTACTGTGCGTACCATCACCACCCAAACCTTCACCC CGATAAATTCTGGCCCTCGAATCGCTGCCAGATGATGACTACATCACCAACCTTCACCC CGACAAAATCTCGGCCCTCGAAATCCGCGAGATGTTACAATCACCAACCGTTAACCC 2627 | 420 |
| 2567 | ACCGAATGACCATCGTGCCTGAACTCTGTGCGTACCATCACCAACCTTCACCC CAGCGAGACCTCGMNGCGTGAAGAGCTACAGTCAGTCAGTCAGTGCGGAACTGGATCAA 282568 | 480 |
| 491 | CAGCGAGAGGCGTGTGAGCTGAAAGCTATAGTCACCGTTGGAACTGGATCAA CAGCGAGAGGCGTGTGAGCTGAAAGCTATAGTCACCGTTGGAACTGGATCAA 282508 | 540 |
| 2507 | CGCGAGAGGCGTGTGAGCTGAAAGCTATAGTCACCGTTGGAACTGGATCAA CGCGAGAGGCGTGTGAGCTGAAAGCTATAGTCACCGTTGGAACTGGATCAA 282448 | 540 |
| 541 | CAGCGTTCGTTGATCGCCCTGATGGCGAGCACTGAGCTGAGCTGAGTCCG CAGCGTTCGTTGATCGCCCTGATGGCGAGCACTGAGCTGAGTCCG 282388 | 600 |
| 601 | CGCCGAGCTTCACTCTGTGAAAMACCCAGCAAAATCTAGTCCCGTGTCAACTGCTGAA CGCCGAGCTTCACTCTGTGAAAMACCCAGCAAAATCTAGTCCCGTGTCAACTGCTGAA 282338 | 660 |
| 6447 | CAGCGTTCGTTGATCGCCCTGATGGCGAGCACTGAGCTGAGTCCG CAGCGTTCGTTGATCGCCCTGATGGCGAGCACTGAGCTGAGTCCG 282388 | 660 |
| 5387 | CGCGATCTGACTCTTGCGAAACCTCGCAAAATCTAGTCCCGTGTCAACTGCTGAA CGCGATCTGACTCTTGCGAAACCTCGCAAAATCTAGTCCCGTGTCAACTGCTGAA 282338 | 660 |
| 661 | GAAATGACCGCCGCTGAGCTGAAACCGACGACGCGTACTGTAGACCGACGATGCG GAAATGACCGCCGCTGAGCTGAAACCGACGACGCGTACTGTAGACCGACGATGCG 720 | 720 |
| 1327 | AAAATGACCGCCGCTGAGCTGAAACCGACGACGCGTACTGTAGACCGACGATGCG AAAATGACCGCCGCTGAGCTGAAACCGACGACGCGTACTGTAGACCGACGATGCG 282268 | 720 |
| 721 | TATTCGTTAACATTGCGATCTAGCCGATAGCCGAAATCATCSCACCATTCGCG TATTCGTTAACATTGCGATCTAGCCGATAGCCGAAATCATCSCACCATTCGCG 780 | 780 |
| 1267 | TATTCGTTAACATTGCGATCTAGCCGATAGCCGAAATCATCSCACCATTCGCG TATTCGTTAACATTGCGATCTAGCCGATAGCCGAAATCATCSCACCATTCGCG 282208 | 780 |
| 781 | TGAGGTTATCGCTTCAGGTAATCTGAGAATTAAT TGAGGTTATCGCTTCAGGTAATCTGAGAATTAAT 818 | 818 |
| 207 | CGAGGTTATCGCTTCAGGTAATCTGAGAATTAAT CGAGGTTATCGCTTCAGGTAATCTGAGAATTAAT 282170 | 282170 |
| AE008916 | Salmonella typhimurium LT2, section 220 of the complete genome. | 13852 bp DNA linear BCT 23-APR-2003 |
| AE008916 | AE008916 AE008916 GI:16423153 | 13852 bp DNA linear BCT 23-APR-2003 |
| AE008916.1 | AE008916.1 GI:16423153 | 13852 bp DNA linear BCT 23-APR-2003 |
| · | Salmonella typhimurium LT2 | |
| | Salmonella typhimurium LT2 | |
| | Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonellae. | |
| 1 | (bases 1 to 1852) | |
| McClelland, M., Sanderson, K. E., Spieth, J., Clifton, S. M., Lareille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, R., Grawal, N., Mulvane, B., Ryan, E., Sun, H., Flores, L., Miller, W., Stoenkeling, T., Nhan, M., Waterston, R. and Wilson, R.K. | | |
| COMMENT | Complete genome sequence of <i>Salmonella enterica</i> serovar Typhimurium LT2 | |
| Nature 413 (6858), 852-856 (2001) | | |
| 21534948 | | |
| 1.677609 | | |
| 2 | (bases 1 to 13852) | |
| | The <i>Salmonella typhimurium</i> Genome Sequencing Project | |
| | Direct Submission | |
| | Submitted (29-MAR-2001) Genome Sequencing Center, Department of Genetics, Washington University School of Medicine, 4444 Forest Park Boulevard, St. Louis, MO 63108, USA | |
| | COMMENT | |
| | Supported by NIH grant 5U 01 AI43283 | |
| | Coding sequences below are predicted from manually evaluated computer analysis, using similarity information and the programs: Glimmer; http://www.tigr.org/soft/blast/glimmer/glimmer.html and Genemark; http://opall.biology.gatech.edu/GenemeMark/ | |
| | EC numbers were directly provided by Junko Yabuzaki and the Kyoto Encyclopedia of Genes and Genomes; http://www.genome.ad.jp/kegg/ , and Pedro Romero and Peter Karp at Ecocyc; | |

<http://ecocyc.PangeaSystems.com/ecocyc/>

The analyses of ribosome binding sites and promoter binding sites were kindly provided by Heladio Salgado, Julio Collado-Vides and RegulonDB; kinich.cifn.unam.mx:8850/db/regulondb_intro.frameset

http://kinich.cifn.unam.mx:8850/db/regulondb_intro.frameset

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistries or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one m13 subclone.

FEATURES Source

Location/Qualifiers

1. .13852 /organism="Salmonella typhimurium LT2" /mol_type="Genomic DNA" /strain="LT2" ATCC 1412; ATCC 700720" /db_xref="ATCC:700720" /db_xref="Taxon:99287" /notes="LT2" /gene="creA" /gene="creA" /notes="synonym: STM4587" /-35_signal 90..98 /gene="creA" /note="putative -35 signal for creA; RegulonDB:STM1H004682" 117..125 /gene="creA" /note="putative RBS for creA; RegulonDB:STM1H004447" 167..640 /note="putative periplasmic protein" /protein_id="AAU23402.1" /db_xref="GI:16423154" /translation="MVKYKSLVTLFSIILLGOSARAQGIVSYDVTVFCKMFGDPDKTVEA FDDPDKRNTVCKVSKRTGGKGKGGLSDSAALISCAQSCQVGPPIELSKDKNGKGQE VVFLKRSLSLKSLSLQVRFDEKRNTLAYLAYSDKVDGSKRNAAASAVPMPWRQ" /gene "creB" .1342 /gene="creB" /note="synonym: STM4588" /gene="creB" /note="putative RBS for creB; RegulonDB:STM1H004448" 653..1342 /gene="creB" /note="OmpR family; similar to E. coli catabolic regulation response regulator (AAU7351.1); Blastp bit to AAU7351.1 (229 aa), 83% identity in aa 1 - 229" /codon_start=1 /transl_table=11 /product="regulatory protein in two-component regulatory system with CreC" /protein_id="AAU23403.1" /db_xref="GI:16423155" /translation="MQQQWVLLVEDBSQGIAQTLIYHQEGFTTVELFARGLPALERK QQRDPIVLDGPDISGFELCRQQLERHPALPILLTASDEVRVLIGBIGADYV AKPSPREVGSKRVTLLRKVKEAASPVRVTRGHLNEPAQQAIAWFGTPSLSIYEP LLLKTILSPRSVYRQMLDWDQETFDRTVTHIKLURAKURAINBELSPINT HRGMGYSILRSV" /transl_table="M0QOWWVLLVEDBSQGIAQTLIYHQEGFTTVELFARGLPALERK QQRDPIVLDGPDISGFELCRQQLERHPALPILLTASDEVRVLIGBIGADYV AKPSPREVGSKRVTLLRKVKEAASPVRVTRGHLNEPAQQAIAWFGTPSLSIYEP LLLKTILSPRSVYRQMLDWDQETFDRTVTHIKLURAKURAINBELSPINT 1330..1335 /gene="creB" /note="putative RBS for creB; RegulonDB:STM1H004449" /gene="creC" /EC_number="2.7.3.-" /transl_table=11 /product="sensory catabolite repression sensor kinase for PhoB; alternative catabolite repressor kinase for PhoB; alternative sensor for pho regulon (AAC77352.1); Blastp hit to AAC77352.1 (474 aa), 80% identity in aa 1 - 474" /codon_start=1 /transl_table=11 /product="sensory kinase (alternative) in two-component regulatory system with CreB (or alternatively PhoB)" /protein_id="AAU23404.1" /db_xref="GI:16423156" /translation="MRIGRLULGFLIVVAVNVLSTFVQLEIKGVRRATEGTLID TATIQLALADDLUSGNPTGQCLAKAFQAOQHPRRANISIGIVKVNNEYVLYWTQG KVLFESENKAQVGDYSRWNDWMLWLTURGOXGARSTKAADPNSITMYVAPITISGRI IGVTVGPKHMANVTTTASVNLGIALGAGGMMWNTSRSLTYAD SVTENRPLA.PALGSESELRLQIAQLESMRKLGKNEIEQVYVALTHLSPLAING AAEIUREGPADWVURFTENLDAONTRMOLVETLROAILENRDIPLAPEVADEF TLSEARSISQIAKQKLTIPSLVVADELLAQLQAGVNDNAIDPENGVILS AQPMBEKAIIQDGGCIGIDFALLFDFPYSIPRENGKSSGIGLAFYSEARILIN GEVTLCNRPQGGVVLASLTIRHFT" 2812..4173 /gene="creD" /note="synonym: STM4590" /gene="creD" /note="putative RBS for creD; RegulonDB:STM1H004450" 2812..2817 /gene="creD" /note="similar to E. coli tolerance to colicin E2 protein" (AAC77353.1); Blastp hit to AAC77353.1 (450 aa), 73% identity in aa 1 - 450" /codon_start=1 /transl_table=11 /product="tolerance to colicin E2 protein" /protein_id="AAU23405.1" /db_xref="GI:16423157" /translation="MLKSPILWKMITLGGAMILLIPLIMMVWVHTTVERADYRSHVEAA IRQTSQGPQKVWGPVIAVPTPELYVLEKEVOKYKSYFYWFLPESLVEGNQVEA RKGTYQGQWHTMDKAQKFVVARLHEARNRNTLGKPEIWGVGDCRISQVAPQO VNGETUTVEGTGLIPESRECISHIPDPSQATRNTLDSLNLSCTGSFSLSLVESS EMTLASNPWNPFWGDFLCKRERSGSGFQWOTSFRAINTLNGSFQFADACKUDWNLV AFSAVASTPDKQIQTDRATVNLVNLWTFMFTTGQRLHMQWQDNLVWMLVW MFLYLNALBBHIGPMWAAISLVMGALMSVYQIAVLMKWNSTLFLTLLALDGVM WGLLUSEDSDSLLGSGVLUALLGGWMFLTRHLDWVSLSCQORKSLPPVVKDDELRLWK" complement(4231..5324) /gene="stb" /note="synonym: STM4591" complement(4231..5316) /gene="stb" /codon_start=1 /transl_table=11 /product="putative major fimbrial subunit" /protein_id="AAU23406.1" /db_xref="GI:16423158" /translation="MRLYVLAIIIFLAFYASCHSYASCKRSNEGATITTPSQLVVDSH ATTAGEVWLGWGMWNTSEVNDGSDRDKYGFPLVPGSGASNTNTINADGNNTPVF STGSGVGWAKTQTMAGPNVMPDNTYNGDNKTHMAMPAVNEVLVALGPIT SGTAFOSPPARVSRSATSEDSGDLHILYLGQVOLINKAMGRVETAITVLDG SVLGSGPANSSTAGGEQDILITLBCGTAISASLRQAPASCGNPDVSIOSNSAAPS TTGVSVOLQORPAGFQSPRSRPLNOKIDPHTHTTINADGSQTINGGTMNSTLKI SARYVTKATAATVTAGONATVNLITYN" complement(5319..5324) /gene="stb" /note="putative RBS for stb; RegulonDB:STM1H004451" complement(5357..5929) /gene="stb" /note="synonym: STM4592" complement(5357..5914)

| | | | |
|-----------------------|---|-----------|---|
| Qy | 752 ACGCGGAAATCATGCCAACATCACGGTGAAAGTTATGCCCTCTGCCGTGACIACGAAAGTCGATAAATTCTCGC | Db | 121 GCGAATGTTGCGTGTAGTGTCTCGACTCGCCGTGACIACGAAAGTCGATAAATTCTCGC |
| Db | 661 ACCCGGAATCATCGTGTACCATTCACGGTAAAGCTGAGCTACGTTCTGGCGATCTGAG 720 | Qy | 377 CTGCAATCGGTSCAGATGACTACATCACCAAAACCGTCAACCCGCGTGACTGACGATT 436 |
| Oy | 812 GATTA 817 | Db | 181 CTGGATATCGGTGAGATGACTACATCACCAAAACCGTCAACCCGCGTGACTGACGATT 240 |
| Db | 721 GATTA 726 | Qy | 437 CGTCACCCAACTTACTGTCCTACATGAGTCAAGTCAACCGTCAACCCGCGTGACTGACGATT 496 |
| Qy | RESULT 14 | Db | 241 CGTCACCCAACTTACTGTCCTACATGAGTCAAGTCAACCGTCAACCCGCGTGACTGACGATT 300 |
| AF267619 | AF267619 Escherichia coli strain E2348/69 aerobic regulator (arcA) gene, partial cdb. | Qy | 497 AGCTTGAAGCTCAAGTCAATGGTGGAACTGGACATCGACAGTCAGAAAGACGCTGGT |
| LOCUS | AF267619 Escherichia coli strain E2348/69 aerobic regulator (arcA) gene, partial cdb. | Db | 301 AGGTTGAAGCTACAAGTCAATGGTGGAACTGGACATCGACAGTCAGAAAGACGCTGGT |
| DEFINITION | | Qy | 557 GGCCCTGATGGCAGCAGTACAAGCTGCAGCTGCAGCTGCAGTTCGGCGCATGCTCACCTC |
| ACCESSION | AF267619 | Db | 361 GGCCCTGATGGCAGCAGTACAAGCTGCAGCTGCAGTTCGGCGCATGCTCACCTC |
| VERSION | AF267619.1 GI:9664513 | Qy | 617 TGTGAAACCCAGCAAAATTCTAGTCCTGTCGTAAGCTGACGTCGGCATCGCGTATGTTGTC |
| KEYWORDS | . | Db | 421 TGTCGAAACCCAGCAAAATTCTAGTCCTGTCGTAAGCTGACGTCGGCATCGCGTATGTTGTC |
| SOURCE | Escherichia coli | Qy | 481 GAGCTGAAACCCAGCAAGCAGACGGTACTGTGACGTCAGCTGCAGTTCGGCATCGCGT |
| ORGANISM | Escherichia coli | Db | 677 GAGCTGAAACCCAGCAAAATTCTAGTCCTGTCGTAAGCTGACGTCAGCTGCAGTTC |
| REFERENCE | 1 (bases 1 to 564) | Qy | 737 GAATCTACGCCGATACCCGGAA 760 |
| AUTHORS | Reid,S.D., Herbelin,C.J., Bumbaugh,A.C., Selander,R.K. and Whittam,T.S. | Db | 541 GAATCTACGCCGATACCCGGAA 564 |
| TITLE | Parallel evolution of virulence in pathogenic Escherichia coli | RECORDS | RESULT 15 |
| JOURNAL | Nature 406 (6791), 64-67 (2000) | FEATURES | AF267620 |
| MDLINE | | source | AF267620 AF267620 Escherichia coli strain DEClA aerobic regulator (arcA) gene, partial cdb. |
| PUBMED | 10894541 | ACCESSION | AF267620.1 GI:9664515 |
| REFERENCE | 2 (bases 1 to 564) | VERSION | |
| AUTHORS | Reid,S.D., Herbelin,C.J., Bumbaugh,A.C., Selander,R.K. and Whittam,T.S. | KEYWORDS | |
| TITLE | Parallel evolution of virulence in pathogenic Escherichia coli | SOURCE | Escherichia coli |
| JOURNAL | Nature 406 (6791), 64-67 (2000) | ORGANISM | Escherichia coli |
| MDLINE | 20351039 | REFERENCE | 1 (bases 1 to 564) |
| PUBMED | 10894541 | AUTHORS | Reid,S.D., Herbelin,C.J., Bumbaugh,A.C., Selander,R.K. and Whittam,T.S. |
| REFERENCE | 2 (bases 1 to 564) | TITLE | Parallel evolution of virulence in pathogenic Escherichia coli |
| AUTHORS | Reid,S.D., Herbelin,C.J., Bumbaugh,A.C., Selander,R.K. and Whittam,T.S. | JOURNAL | Nature 406 (6791), 64-67 (2000) |
| TRANSLATION | "ATGGAEMQQILSBYDINWIMDINLPKRNGLIARELREQANVA LMFLTERDNEDVKEKLLIGEADDYTKPNFRELTRARNLISRTMLNGTSEERSV ESYKFNGWELDINSLSLGPOEVKLPRSFAMLHFCENPGKLOSRAELKMTGR | FEATURES | Submitted (12-MAY-2000) Biology, Pennsylvania State University, 208 |
| ORIGIN | | source | Mueller Laboratory, University Park, PA 16802, USA |
| Query Match | 60.7%; Score 562.4; DB 1; length 564; | ACCESSION | Location/Qualifiers |
| Best Local Similarity | 99.8%; Pred. No. 5.2e-156; | VERSION | |
| Matches | 563; Conservative 0; Mismatches 1; Indels 0; Gaps 0; | KEYWORDS | |
| Oy | 197 GCGACAGATGCCGGAAATGATCAGTCTCTGATATGATCACCAACCTGGTAC 256 | SOURCE | Escherichia coli |
| Db | 1 GCGACAGATGCCGGAAATGATCAGTCTCTGATATGATCACCAACCTGGTAC 60 | ORGANISM | Enterobacteriaceae; Escherichia. |
| Qy | 257 ATGGATATGCACTGCCGGTAAGAACGGTCTCTGTAGCGCTGACTGCCGACAG 316 | REFERENCE | 1 (bases 1 to 564) |
| Db | 61 ATGGATATGCACTGCCGGTAAGAACGGTCTCTGTAGCGCTGACTGCCGACAG 120 | AUTHORS | Reid,S.D., Herbelin,C.J., Bumbaugh,A.C., Selander,R.K. and Whittam,T.S. |
| Qy | 317 GCGATGTTGGTGTAGTCTCTGACTGCCGACAGAAGTCGATAAATCTCGC 376 | TITLE | Parallel evolution of virulence in pathogenic Escherichia coli |
| ORIGIN | | JOURNAL | Nature 406 (6791), 64-67 (2000) |
| Query Match | 60.7%; Score 562.4; DB 1; length 564; | FEATURES | Submitted (12-MAY-2000) Biology, Pennsylvania State University, 208 |
| Best Local Similarity | 99.8%; Pred. No. 5.2e-156; | source | Mueller Laboratory, University Park, PA 16802, USA |
| Matches | 563; Conservative 0; Mismatches 1; Indels 0; Gaps 0; | ACCESSION | Location/Qualifiers |
| Oy | 197 GCGACAGATGCCGGAAATGATCAGTCTCTGATATGATCACCAACCTGGTAC 256 | VERSION | |
| Db | 1 GCGACAGATGCCGGAAATGATCAGTCTCTGATATGATCACCAACCTGGTAC 60 | KEYWORDS | |
| Qy | 257 ATGGATATGCACTGCCGGTAAGAACGGTCTCTGTAGCGCTGACTGCCGACAG 316 | SOURCE | Escherichia coli |
| Db | 61 ATGGATATGCACTGCCGGTAAGAACGGTCTCTGTAGCGCTGACTGCCGACAG 120 | ORGANISM | Enterobacteriaceae; Escherichia. |
| Qy | 317 GCGATGTTGGTGTAGTCTCTGACTGCCGACAGAAGTCGATAAATCTCGC 376 | REFERENCE | 1 (bases 1 to 564) |
| ORIGIN | | AUTHORS | Reid,S.D., Herbelin,C.J., Bumbaugh,A.C., Selander,R.K. and Whittam,T.S. |
| Query Match | 60.7%; Score 562.4; DB 1; length 564; | TITLE | Parallel evolution of virulence in pathogenic Escherichia coli |
| Best Local Similarity | 99.8%; Pred. No. 5.2e-156; | JOURNAL | Nature 406 (6791), 64-67 (2000) |
| Matches | 563; Conservative 0; Mismatches 1; Indels 0; Gaps 0; | FEATURES | Submitted (12-MAY-2000) Biology, Pennsylvania State University, 208 |
| Oy | 197 GCGACAGATGCCGGAAATGATCAGTCTCTGATATGATCACCAACCTGGTAC 256 | source | Mueller Laboratory, University Park, PA 16802, USA |
| Db | 1 GCGACAGATGCCGGAAATGATCAGTCTCTGATATGATCACCAACCTGGTAC 60 | ACCESSION | Location/Qualifiers |
| Qy | 257 ATGGATATGCACTGCCGGTAAGAACGGTCTCTGTAGCGCTGACTGCCGACAG 316 | VERSION | |
| Db | 61 ATGGATATGCACTGCCGGTAAGAACGGTCTCTGTAGCGCTGACTGCCGACAG 120 | KEYWORDS | |
| Qy | 317 GCGATGTTGGTGTAGTCTCTGACTGCCGACAGAAGTCGATAAATCTCGC 376 | SOURCE | Escherichia coli |
| ORIGIN | | ORGANISM | Enterobacteriaceae; Escherichia. |

/db xref="GI:9664516"
 /translation="ATDGAEMMQLSYDINIVMDINLRKGNLARELREBONA
 LMFIIGRONEVDKILGIGIAGADYITKPFPNPFELTRARNLSRTMNICVSEERSV
 ESYKENGHLDINSRSLLGPDRQYKUPRSFFRAMLFCEPFGKIQSABLLKOMTGR
 ELKPHDRTVDTIRRKFESTPDITE"

ORIGIN

| Query | Match | 60.7% | Score | 562.4; | DB | 1; | Length | 564; |
|-------|---------|---|--------------|--------|---------|------|--------------|-----------|
| | Bast | Local | Similarity | 99.8% | Matches | 563; | Conservative | 5.2e-156; |
| | Matches | 563; | Conservative | 0; | Pred. | No. | Mismatches | 1; |
| | | | | | Indels | 0; | Gaps | 0; |
| QY | 197 | GCGACAGAGGGCGGGAAATGCCATCAGATCCCTCTGAAATGACATCACTGGTGCATC | 256 | | | | | |
| Db | 1 | GCGACAGAGATGGCGGGAAATGCCATCAGATCACTGGTGCATC | | | | | | |
| QY | 257 | ATGCAATGCAATCTGCCGGTAAGAACGGCTCTGTAGCCGGTAACCTGGTGCATC | 60 | | | | | |
| Db | 61 | ATGGATATACTTGCCGGTAAAGAAGGTTCTCTGTAGCCGGTAACCTGGTGCATC | | | | | | |
| QY | 317 | GCGATGTTGGTGTATGTTCTGACTGGCCCTGACAGAGTCATAATTCTCGCC | 376 | | | | | |
| Db | 121 | GCGATGTTGGTGTATGTTCTGACTGGCCCTGACAGAGTCATAATTCTCGCC | | | | | | |
| QY | 377 | CTCGAAATGGTGTACAGATGACTACATCACAAACCTTCACCCGGTGAATGAGATT | 436 | | | | | |
| Db | 181 | CTCGAAATGGTGTACAGATGACTACATCACAAACCTTCACCCGGTGAATGAGATT | | | | | | |
| QY | 437 | CCTGACGGAACTACTGTTCCGTACATCATCGGGTCTGCAAGCGAGACGTGT | 496 | | | | | |
| Db | 241 | CCTGACGGAACTCTGTCGTCGGTACATCATCGGGTCTGCAAGCGAGACGTGT | | | | | | |
| QY | 497 | ACGGTGAAGACTACAGTCATGCTGCAACTGGCACTGGACTCAACGGCGTGTGATC | 556 | | | | | |
| Db | 301 | ACGGTGAAGACTACAGTCATGCTGCAACTGGCACTGGACTCAACGGCGTGTGATC | | | | | | |
| QY | 557 | GGCCCTGATGGGCGCACTGCAAGTCAGGTGCGGACTTCCGGCATGGCTCACTTC | 616 | | | | | |
| Db | 361 | GGCCCTGATGGGCGCACTGCAAGTCAGGTGCGGACTTCCGGCATGGCTCACTTC | | | | | | |
| QY | 617 | TGTGAAACCCAGGAATTCAGTCGCCCTGCTGACTGCTGAGAAATGACCGGCGT | 676 | | | | | |
| Db | 421 | TGTGAAACCCAGGAATTCAGTCGCCCTGCTGACTGCTGAGAAATGACCGGCGT | | | | | | |
| QY | 677 | GAGCTGAAACGGAGACCTACTGTAGAGCTGAGATCCGGTATTCGAAACCTTC | 736 | | | | | |
| Db | 481 | GAGCTGAAACGGAGACCTACTGTAGAGCTGAGATCCGGTATTCGAAACCTTC | | | | | | |
| QY | 737 | GAGCTACCGGATAACCGGAA | 760 | | | | | |
| Db | 541 | GAGCTACCGGATAACCGGAA | 564 | | | | | |

Search completed: January 22, 2005, 13:53:24
 Job time : 4273 secs

CC aerobic respiration control) protein does not normally function. The CC disruption to ArcA protein function is preferably due to disruption of a CC chromosomal arca gene. The bacterium and method are useful in producing a CC target substance by fermentation. The target substance is preferably an L CC -amino acid, especially L-lysine or L-glutamic acid. Production of the CC present sequence is that of a PCR primer which was used for amplification CC of the P *anaerobius* *arcA* gene in the exemplification of the invention.

XX

Sequence 927 BP; 236 A; 228 C; 225 G; 238 T; 0 U; 0 Other;

Query Match 100.0%; Score 927; DB 12; Length 927;
Best Local Similarity 100.0%; Pred. No. 1.1e-297; Mismatches 0; Indels 0; Gaps 0;

Matches 927; Conservative 0; MisMatches 0; Indels 0; Gaps 0;

QY 1 GTCATGTTACCGGATCATGTTAATTTCGACATCATTGGAGGTCTGGACTTGT

Db 1 GTCATGTTACCGGATCATGTTAATTTCGACATCATTGGAGGTCTGGACTTGT

QY 61 ACTTCCTGTTGATGTTAGTGGCAATTAGTAGCAACATCGAGCCCCACATTCT

Db 61 ACTTCCTGTTGATGTTAGTGGCAATTAGTAGCAACATCGAGCCCCACATTCT

QY 121 TATCGTTGAAAGCAGAGTTGTTAACAGCAACCTTGAAAGTTTGAGGGAGG

Db 121 TATCGTTGAAAGCAGAGTTGTTAACAGCAACCTTGAAAGTTTGAGGGAGG

QY 181 CTATCGATGTTTTCGAAAGCAGATGGCCGGAAATGCATCAGATCCCTCTGATATGA

Db 181 CTATCGATGTTTTCGAAAGCAGATGGCCGGAAATGCATCAGATCCCTCTGATATGA

QY 241 CATCAACCTGGATCATGGATACTGCGCGTAAAGCAGGCTCTTGTGTTGCG

Db 241 CATCAACCTGGATCATGGATACTGCGCGTAAAGCAGGCTCTTGTGTTGCG

QY 301 TGAAGTCGGAGAGGGCAATGTGCGGTGAGTTCTCTGACTGGGGTGAAACGAGT

Db 301 TGAAGTCGGAGAGGGCAATGTGCGGTGAGTTCTCTGACTGGGGTGAAACGAGT

QY 361 CGATAAAATTCTCGGCCCTCGAAATCGGGCAATGACTCACCAACGGTCAACC

Db 361 CGATAAAATTCTCGGCCCTCGAAATCGGGCAATGACTCACCAACGGTCAACC

QY 421 GCGTAACGTGACCTCGCACTTGCGGAACTTGCGTCCGTCATGCTGGACTGT

Db 421 GCGTAACGTGACCTCGCACTTGCGGAACTTGCGTCCGTCATGCTGGACTGT

QY 481 CACGGGAGGACGCGTAGCGTGAAGCTGTCATGTCATGTCATGTCATGTCATCAA

Db 481 CACGGGAGGACGCGTAGCGTGAAGCTGTCATGTCATGTCATGTCATGTCATCAA

QY 541 CACCGGTCTGTCATCGGCCCTGATCGCGACGAGTACAAGGTGGCGCGAGGACTTCG

Db 541 CACCGGTCTGTCATCGGCCCTGATCGCGACGAGTACAAGGTGGCGCGAGGACTTCG

QY 601 CGCGATGTCATCTCGTGAAGCTGTCATGTCATGTCATGTCATGTCATGTCATGAA

Db 601 CGCGATGTCATCTCGTGAAGCTGTCATGTCATGTCATGTCATGTCATGTCATGAA

QY 660 601 CGCGATGTCATCTCGTGAAGCTGTCATGTCATGTCATGTCATGTCATGTCATGAA

Db 660 601 CGCGATGTCATCTCGTGAAGCTGTCATGTCATGTCATGTCATGTCATGTCATGAA

QY 661 GAAATGACCCCGCGAGCTGAAACCGAGGACCTACTGTAGCTGAGATCCGCG

Db 661 GAAATGACCCCGCGAGCTGAAACCGAGGACCTACTGTAGCTGAGATCCGCG

QY 720 661 GAAATGACCCCGCGAGCTGAAACCGAGGACCTACTGTAGCTGAGATCCGCG

Db 720 661 GAAATGACCCCGCGAGCTGAAACCGAGGACCTACTGTAGCTGAGATCCGCG

QY 780 720 721 TATCGTAAACATTGAACTACCGCCGATACGCCGAAATCTACCGTACCGTACCG

Db 780 720 721 TATCGTAAACATTGAACTACCGCCGATACGCCGAAATCTACCGTACCGTACCG

QY 781 TGAAGCTTATGCTCTGCGCTGCTGAGGAGATATCGCTTACCGTACCGTACCG

Db 781 TGAAGCTTATGCTCTGCGCTGCTGAGGAGATATCGCTTACCGTACCGTACCG

QY 840 781 841 AACGGGCTTTAGGCCGTTTATTTCACCTTATTCCAGTACCTCATC

Db 840 841 AACGGGCTTTAGGCCGTTTATTTCACCTTATTCCAGTACCTCATC

QY 900 840 900 Sequence 1732 BP; 439 A; 391 C; 417 G; 477 T; 0 U; 8 Other;

RESULT 2

ID ABL90421/C

XX

ID ABL90421

XX

AC

ABL90421;

XX

DT

XX

DE

Human polynucleotide SEQ ID NO 983.

XX

CYTOSATIC: immunosuppressive; nootropic; neuroprotective; antiviral;

KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;

KW vulnerable; anticonvulsant; antibacterial; antifungal; antiparasitic;

KW cardiotonic; gene therapy; cancer; immune disorder; cardiovascular disorder;

XX

OS Homo sapiens.

XX

PN WO200190304-A2.

XX

PR 19-MAY-2000; 2000US-0205515P.

XX

(HUMA-) HUMAN GENOME SCI INC.

XX

PT

Birse CE, Rosen CA;

XX

DR WPI; 2002-122018/16.

XX

P-PSDB; ABB90012.

XX

PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and

PT prevention of neural, immune system, muscular, reproductive,

PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative

PT disorders.

XX

Claim 4; SEQ ID NO 983; 2081PP + Sequence Listing; English.

XX

The invention relates to novel genes (ABB89442-ABB90853) and proteins

CC (ABB89400-ABB90444) useful for preventing, treating or ameliorating

CC medical conditions e.g. by protein or gene therapy. The genes are

CC isolated from a range of human tissues disclosed in the specification.

CC The nucleic acids, proteins, antibodies and (anti)agonists are useful in

CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and

CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,

CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune

CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic

CC anaemia, autoimmunthyroiditis, diabetes mellitus, Crohn's disease,

CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)

CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing

CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)

CC infectious diseases such as viral, bacterial, fungal and parasitic

CC infections. Note: The sequence data for this patent did not form part of

CC the printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

Sequence 1732 BP; 439 A; 391 C; 417 G; 477 T; 0 U; 8 Other;

XX

Query Match 99.9%; Score 925.8; DB 6; Length 1732;

Best Local Similarity 99.7%; Pred. No. 3.9e-297;

Matches 924; Conservative 3; MisMatches 0; Indels 0; Gaps 0;

QY 1 GTCATGTTACCGGATCATGTTAATTTCGAGTACGTCAGGCTGGAGCTTGT

| | | | | | |
|---|------------------------|---|-------|----|---|
| Db | 1313 | GTCATGTTAGCCGATCAGTTATTGCAAGCATCAGGAGGTAGGGACTTGT | 1254 | KW | Recombinant expression vector; transcription regulatory element; |
| Oy | 61 | ACTCTCTGTTCTGATTAAGTGCAATTAGGTAGCACAAACATCAGACCCCGCACATCT | 120 | XX | Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds. |
| Db | 1253 | ACTCTCTGTTCTGATTAAGGTAGCTGGCATTTGGCAATTCAGAACACATCT | i 194 | XX | Klebsiella pneumoniae. |
| Oy | 121 | TATCGTTGAGAGGAGTTGGAACACGCCAACAGTTGAAAGTATTTCGAAGAGGAGG | 180 | XX | |
| Db | 1193 | TATCGTTGAGAGGAGTTGGAACACGCCAACAGTTGAAAGTATTTCGAAGAGGAGG | 1134 | PN | US610936-B1. |
| Oy | 191 | CTATGATGTTTCAAGCACAAGATGGCGGGAAATGCATCAGATCCCTCTGAAATGA | 240 | XX | |
| Db | 1133 | CTATGATGTTTCAAGCACAAGATGGCGGGAAATGCATCAGATCCCTCTGAAATGA | 1074 | PR | 29-JAN-1999; 990US-0117747P. |
| Oy | 301 | TGAACTGGGGACAGAGCATGGGGGAAATGATCAGATCCCTCTGAAATGA | 360 | PA | (GENO-) GENOME THERAPEUTICS CORP. |
| Db | 241 | CATCAACCTGTTGATCATCGATAATCTGCGCGTAGAAGACCGCTCTCGTTAGCG | 300 | XX | |
| Oy | 1073 | CATCAACCTGTTGATCATCGATAATCTGCGCGTAGAAGACCGCTCTCGTTAGCG | 1014 | PR | 27-JAN-2000; 2000US-00489039. |
| Oy | 361 | CGATAAAATTCTCGCCCTGAAATGGTCCAGTACTGGCTGAGATGACTACATCACC | 420 | XX | |
| Db | 953 | CGATAAAATTCTCGCCCTGAAATGGTCCAGTACTGGCTGAGATGACTACATCACC | 894 | DR | WPI; 2003-893346/82. |
| Oy | 421 | GCGGAACTGAGGATCGGATCGGAGCCACCTACTGCTCCGATCCAGTAATCTGGTACTGT | 480 | XX | P-PSDB; AB082466. |
| Db | 893 | GCGGAACTGAGGATCGGATCGGAGCCACCTACTGCTCCGATCCAGTAATCTGGTACTGT | 834 | XX | |
| Oy | 481 | CAGCGGAAGAAGCTGCTGAGCTGAAAGCTAACAGTCAATGTTGGAAACTGACATCAA | 540 | PT | New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for preparing a vaccine composition against Klebsiella pneumoniae. This sequence encodes a Klebsiella pneumoniae polypeptide of the invention. |
| Db | 833 | CAGCGGAAGAAGCTGCTGAGCTGAAAGCTAACAGTCAATGTTGGAAACTGACATCAA | 774 | PT | Preparation; SEQ ID NO 1812; 932pp; English. |
| Oy | 541 | CAGCGGTCTGATCGCCCTGATGGCGGAGCACTACAGCTCGGGCGAGCGAGTCCG | 600 | PS | Disclosure; SEQ ID NO 1812; 932pp; English. |
| Db | 773 | CAGCGGTCTGATGGCGGAGCACTACAGCTCGGGCGAGCGAGTCCG | 714 | XX | |
| Oy | 601 | CGCGATGCTTACATCTGTAACCCAGGCGAAATTAGTGGCCCGTGTGAACTGCTAA | 660 | CC | The invention describes a new isolated nucleic acid encoding a Klebsiella pneumoniae polypeptide. Also described are: a recombinant expression vector comprising the nucleic acid operably linked to a transcription regulatory element; and a cell comprising the recombinant expression vector. The nucleic acid is useful for preparing a vaccine composition against Klebsiella pneumoniae. This sequence encodes a Klebsiella pneumoniae polypeptide of the invention. |
| Db | 713 | CGCGATGCTTACATCTGTAACCCAGGCGAAATTAGTGGCCCGTGTGAACTGCTAA | 654 | CC | |
| Oy | 661 | GAAATGACCGCCGCTGAAACCCAGGCGACTTGAGCTGAGATCCG | 720 | CC | |
| Db | 653 | GAAATGACCGCCGCTGAAACCCAGGCGACTTGAGCTGAGATCCG | 594 | CC | |
| Oy | 721 | TATCGTAAACATTGAACTACGCCGAGATGCCGAATCATGCCAACCTTCAGG | 780 | Db | 1 GTAGCAACATCGACGACCCGCAATCTTATCGTGTGAGAGGAGTGTAAACCGAAC |
| Db | 593 | TATCGTAAACATTGAACTACGCCGAGATGCCGAATCATGCCAACCTTCAGG | 534 | Oy | 152 ACGTGAAAGTATTGAGGGAGCTATGATTTGAGGACAGATGGCG |
| Oy | 781 | TGAGGTATCGCTCTGGGTATCTGGAGAGTTATCGCTTACACCGTCAAAAAA | 840 | Db | 61 ACGTTAAAGTATTTCGAAGCAGAGTTACGATGAGCCATGGCG |
| Db | 533 | TGAGGTATCGCTCTGGGTATCTGGAGAGTTATCGCTTACACCGTCAAAAAA | 474 | Oy | 212 GAAATGCTCAGATCTCTCTGATATGACATCACCTGCTGATGATCAATCTG |
| Oy | 841 | AACGGCGCTTCTAGCGCTTTATTTCAACCTATTCGAGATGACTCATC | 900 | Db | 121 GAAATGCTCAGATCTCTCTGAAAGATATCACCTGGTGTACGATCACCTG |
| Db | 473 | AACGGCGCTTCTAGCGCTTTATTTCAACCTATTCGAGATGACTCATC | 414 | Oy | 272 CCGGGTAGAACGGCTCTCTGAGCTGGAGCAGGGGAATGCGTG |
| Oy | 901 | GTCGGTTACTCTTACTGGCTT | 927 | Db | 181 CCGGGTAGAACGGCTCTCTGAGCTGGAGCAGGGAGTGGCTG |
| Db | 413 | GTCGGTTACTCTTACTGGCTT | 387 | Oy | 332 ATGTTCTGACTGGCTGACACGAGTCATAATTCTGGCTGAATCGGCCA |
| RESULT 3 | | | | Db | 391 301 GAGGACTATACATACCAACCGTTAACCGCGTGAAGTACGATCCCGCGCAACTG |
| ACH9017 | | | | Oy | 452 CTGTCCTGACTGAGCTGGCTACTCTGAGCACGAGAACGCGTGTGATCGCCCTGATGGCG |
| ID ACH96017 | standard; DNA; 726 BP. | | | Db | 361 CTCTCCCTTACATGACCTCGTGTGAGCTGGCGACCGTGGAAAGTAC |
| XX | | | | Oy | 512 ACTCTCTGTTGAGCTGGCTACTCTGAGCACGAGAACGCGTGTGATCGCCCTGATGGCG |
| DT 29-JUL-2004 (first entry) | | | | Db | 421 AACGTCACGGCTGGACTCGATATCACAGCGTGTGAGCGACGCGAA |
| XX | | | | Oy | 572 CAGTACAGCTGGCGAGCGAGTTCGGCGCATGTTACTCTGAAACCGAGC |
| DE Klebsiella pneumoniae polynucleotide seqid 1812. | | | | | 631 |

| WP | Fragment Name | Begin | End | DE |
|-----|--|--------|--------|-----|
| WP | ACF65386_0 | 1 | 11000 | Db |
| WP | ACF65386_1 | 100001 | 210000 | Db |
| WP | ACF65386_2 | 200001 | 310000 | Qy |
| WP | ACF65386_3 | 300001 | 410000 | Qy |
| WP | ACF65386_4 | 400001 | 510000 | Db |
| WP | ACF65386_5 | 500001 | 610000 | Qy |
| WP | ACF65386_6 | 600001 | 700779 | Db |
| ID | ACF65386 standard; DNA; 700779 BP. | | | XX |
| AC | ACF65386; | | | XX |
| DT | 20-NOV-2003 (first entry) | | | XX |
| DE | Photorhabdus luminescens nucleotide sequence #39. | | | DE |
| XX | | | | XX |
| KW | Antibacterial; fungicide; insecticide; polymorphism; genetic analysis; detection; food; gene expression; plant; animal; microorganism; toxin; antibiotic; biopesticide; virulence factor; disease model; plague; whooping cough; gene; ds. | | | XX |
| OS | Photorhabdus luminescens. | | | OS |
| XX | W020294867-A2. | | | XX |
| PD | 28-NOV-2002. | | | PD |
| PR | 07-FEB-2002; 2002WO-1B003040. | | | PR |
| XX | 07-FEB-2001; 2001FR-00001659. | | | XX |
| PA | (INSP) INST PASTER. (CNRS) CENT NAT RECH SCI. | | | PA |
| XX | Duchaud E, Taurit S, Glaser P, Frangeul L, Kunst F, Danchin A; | | | XX |
| PI | Buchrieser C; | | | PI |
| XX | DR. WPI; 2003-148459/14. | | | XX |
| PT | Genomic sequence of Photorhabdus luminescens and encoded polypeptides, useful e.g. as therapeutic antimicrobials and agricultural pesticides. | | | PT |
| XX | Claim 1; SEQ ID NO 39; 1205pp; French. | | | XX |
| CC | The invention relates to the isolation of genes and their encoded proteins from Photorhabdus luminescens. The isolated sequences are sources of probes and primers for detecting the genome of <i>P. luminescens</i> and related species; to study polymorphisms; for gene analysis and for detection/amplicification of the genes. Antibodies (Ab) raised against the polypeptides encoded by the genes are used for detection/identification of <i>P. luminescens</i> , e.g. in foods. The genes, proteins, Ab and cells that carry a gene-containing vector are used to select compounds that modulate, regulate, induce or inhibit expression of the genes in plants, animals or microorganisms other than <i>P. luminescens</i> and are able to alter response or sensitivity to toxins and antibiotics produced by <i>P. luminescens</i> . Cells transformed to express the genes are useful for recombinant production of the proteins, particularly toxins and antibiotics useful as insecticides, bactericides and fungicides. The genes, proteins, vectors containing the genes and Ab are also useful therapeutically (to treat microbial infection by bacteria or fungi that are sensitive to <i>P. luminescens</i> -encoded toxins or antibiotics) and as biopesticides. Other uses of the genes and the proteins are as virulence factors and for identifying targets of human diseases for which <i>P. luminescens</i> is a model (particularly plague and whooping cough). This sequence represents one of the isolated <i>P. luminescens</i> genes. | | | CC |
| CC | result 6 | | | CC |
| CC | ADJ38390 | | | CC |
| CC | ID ADJ38390 standard; DNA; 759 BP. | | | CC |
| CC | XX | | | CC |
| CC | AC ADJ38390; | | | CC |
| CC | XX | | | CC |
| CC | DT 06-MAY-2004 (first entry) | | | CC |
| CC | DE | | | CC |
| CC | Pantoea ananatis arcA (aerobic respiration control) gene SeqId9. | | | CC |
| SQ | Gamma-proteobacterium; arcA; aerobic respiration control; arcA protein function; chromosomal arcA gene; fermentation; L-amino acid; L-lysine; L-glutamic acid; gene; ds. | | | SQ |
| XX | Pantoea ananatis. | | | XX |
| Key | Location/Qualifiers | | | Key |

PT FT CDS 41. .757
 PT /!*arg= a
 PT /product= "Pantoea ananatis arca protein"
 XX EP1382686-A1.
 XX PD 21-JAN-2004.
 XX PR 11-JUL-2003; 2003EP-00015911.
 XX PR 12-JUL-2002; 2002EP-00203764.
 XX PA (AIN) AJINOMOTO CO INC.
 XX PT Ishikawa Y, Imaiizumi A, Matsui K, Kojima H;
 XX WPI; 2004-12-2975/19.
 DR P-PSDB; ADJ38391.
 XX PT New bacterium useful for producing a target substance (especially L-lysine, L-glutamic acid or another L-amino acid) is modified so that the Arca protein does not normally function.
 PS Example 4; SEQ ID NO 19; 38pp; English.

CC This invention relates to a novel gamma-proteobacterium able to produce a target substance such as an amino acid which is modified so that an Arca (aerobic respiration control) protein does not normally function. The disruption to Arca protein function is preferably due to disruption of a chromosomal arca gene. The bacterium and method are useful in producing a target substance fermentation. The target substance is preferably an L-amino acid, especially L-lysine or L-glutamic acid. Production of the target substance using the modified bacterium is more efficient. The present sequence is that of the P ananatis gene which encodes the Arca protein and which was used in the exemplification of the invention.
 XX Sequence 759 BP; 205 A; 192 C; 183 G; 179 T; 0 U; 0 Other;
 CC Best Local Similarity 54.9%; Score 508.6; DB 12; Length 759;
 CC Matches 607; Conservative 0; Mismatches 144; Indels 2; Gaps 1;
 CC Qy 69 TTGCGATTGTGGC-AATTTAGTGCACACATCGAGACCCGACATTCCTATCGT 126
 CC Db 7 TTTCATTGTTGTCGAAATTAGTAGCTAACATCGAGACCCGACATTCCTATCGT 126
 CC Qy 127 TGAAGACGGTTGTAACGCCAACACCTTGAAAGATTTGCGGGAGGTATGA 186
 CC Db 67 TGAAGACGAACTGTCATCGCAATACCTCCAAAGATTGCGGAGGTATGT 125
 CC Qy 187 TGTTCGAGGAGCACAGATGCCGAATCATCGATCTCTGAAATGACATCAA 246
 CC Db 127 CGTGTACGAGGCCACCGAATGTCAGATGCGACATGCGACATGATGCTAA 186
 CC Qy 247 CCTGGTGTACATGATACTAACCTGGGGAGACGGTTCTGTAGCGCTGAAT 305
 CC Db 187 TCTGGTATATGACATCACTGGGGAAACGCGCTGTACTGGCACTGAAT 246
 CC Qy 307 GGCAGACGAGSGAAATGTTGGTGTAGTCTGACTGGCGTACAAGAATGATAA 366
 CC Db 247 GGCTGAGGAAKCCAAATGCGATTGTTCTGACCGGATAAGAAGTCGATA 306
 CC Qy 367 AATTCCTGGCTGCAATCGCAGATGACATCACCAACCGTCAACCGGCGA 426
 CC Db 307 AACATCTGGGTGGAATTCGTCAGACGACTATCTATAGCGTTAACCCACCGA 366
 CC Qy 427 ACTGAGCATGTTGCAAGCAACCTACTGCGCGTACATGAACTGTGTTACTTCAGCGA 486
 CC Db 367 ATTAATCTTGCTGCAACCTGTCGTCGCGACCATGATTTGCCATTACCATGA 426
 CC Qy 487 AGAACCTCGTGAAGCTAACAGTCAATGCTGTTGCGACATACAGCGC 546
 CC Db 427 AGAGGCTGCGACGGTGAAGCTATAGTCATACTGGCTGGAGCTGGACATACAGCG 486

Query Match 54.9%; Score 508.6; DB 12; Length 759;
 Matches 607; Conservative 0; Mismatches 144; Indels 2; Gaps 1;
 Qy 69 TTGCGATTGTGGC-AATTTAGTGCACACATCGAGACCCGACATTCCTATCGT 126
 Db 7 TTTCATTGTTGTCGAAATTAGTAGCTAACATCGAGACCCGACATTCCTATCGT 126
 Qy 127 TGAAGACGGTTGTAACGCCAACACCTTGAAAGATTTGCGGGAGGTATGA 186
 Db 67 TGAAGACGAACTGTCATCGCAATACCTCCAAAGATTGCGGAGGTATGT 125
 Qy 187 TGTTCGAGGAGCACAGATGCCGAATCATCGATCTCTGAAATGACATCAA 246
 Db 127 CGTGTACGAGGCCACCGAATGTCAGATGCGACATGCGACATGATGCTAA 186
 Qy 247 CCTGGTGTACATGATACTAACCTGGGGAGACGGTTCTGTAGCGCTGAAT 305
 Db 187 TCTGGTATATGACATCACTGGGGAAACGCGCTGTACTGGCACTGAAT 246
 Qy 307 GGCAGACGAGSGAAATGTTGGTGTAGTCTGACTGGCGTACAAGAATGATAA 366
 Db 247 GGCTGAGGAAKCCAAATGCGATTGTTCTGACCGGATAAGAAGTCGATA 306
 Qy 367 AATTCCTGGCTGCAATCGCAGATGACATCACCAACCGTCAACCGGCGA 426
 Db 307 AACATCTGGGTGGAATTCGTCAGACGACTATCTATAGCGTTAACCCACCGA 366
 Qy 427 ACTGAGCATGTTGCAAGCAACCTACTGCGCGTACATGAACTGTGTTACTTCAGCGA 486
 Db 367 ATTAATCTTGCTGCAACCTGTCGTCGCGACCATGATTTGCCATTACCATGA 426
 Qy 487 AGAACCTCGTGAAGCTAACAGTCAATGCTGTTGCGACATACAGCGC 546
 Db 427 AGAGGCTGCGACGGTGAAGCTATAGTCATACTGGCTGGAGCTGGACATACAGCG 486

RESULT 7
 ID ACF6984
 ID ACF6984 standard; DNA; 717 BP.
 XX AC F6844;
 XX DT 20-NOV-2003 (first entry)
 XX DE Photorhabdus luminescens nucleotide sequence #8311.
 XX KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 XX KW detection; food; gene expression; plant; animal; microorganism; toxin;
 XX KW antibiotic; biopesticide; virulence factor; disease model; plague;
 XX KW whooping cough; gene; ds;
 OS Photorhabdus luminescens.
 XX PN WO200294867-A2.
 PD 28-NOV-2002.
 XX P07-FEB-2002; 2002WO-1B003040.
 XX P07-FEB-2001; 2001FR-00001659.
 XX PA (INSP) INST PASTEUR.
 XX (CNRS) CNRS CENT NAT RECH SCI.
 XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 XX Buchrieser C;
 XX DR WPI; 2003-148459/14.
 XX PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 PS Claim 2; SEQ ID NO 8311; 120pp; French.
 XX
 CC The invention relates to the isolation of genes and their encoded
 proteins from Photorhabdus luminescens. The isolated sequences are
 sources of probes and primers for detecting the genome of P. luminescens
 and related species; to study polymorphisms; for gene analysis and for
 detection/amplification of the genes. Antibodies (Ab) raised against the
 polypeptides encoded by the genes are used for detection/identification
 of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 carry a gene-containing vector are used to select compounds that
 modulate, regulate, induce or inhibit expression of the genes in plants,
 animals or microorganisms other than P. luminescens and are able to alter
 response or sensitivity to toxins and antibiotics produced by P.
 luminescens. Cells transformed to express the genes are useful for
 recombinant production of the proteins, particularly toxins and

| | | | | |
|------------------------------------|---|----|--|---|
| CC | antibacterials useful as insecticides, bactericides and fungicides. The genes, proteins, vectors containing the gene and Ab are also useful as therapeutically (to treat microbial infection by bacteria or fungi that are sensitive to <i>P. luminescens</i> -encoded toxins or antibiotics) and as biopesticides. Other uses of the genes and the protein are as virulence factors and for identifying targets of human diseases for which <i>P. luminescens</i> is a model (particularly plague and whooping cough). This sequence represents one of the isolated <i>P. luminescens</i> gene | XX | KW | Proteus mirabilis infection; bacterial infection; antibacterial; immunostimulant; gene; ds. |
| CC | | XX | XX | Proteus mirabilis. |
| CC | | OS | US6605709-B1. | |
| CC | | XX | PD | 12-AUG-2003. |
| XX | Sequence 717 BP; 213 A; 158 C; 158 G; 188 T; 0 U; 0 Other; | XX | PP | 05-APR-2000; 2000US-00543681. |
| Query Match | 53.2%; Score 493; DB 10; Length 717; Best Local Similarity 80.5%; Pred. No. 4.2e-153; Matches 577; Conservative 0; Mismatches 140; Indels 0; Gaps 0; | XX | PR | 09-APR-1999; 99US-012876P. |
| Oy | 101 ATTCAGAGCCCCGACATCTCTATCGTCAAAGCAGAGTGGTAACACGCAACAGTTGAA 60 | XX | PA | (GENO-) GENOME THERAPEUTICS CORP. |
| Db | 1 ATGCAAAACCCCGACATTGTGATGTTGAAAGCAACTTGATCATGGATACTCGCAATAACCTAAA 60 | XX | PI | Bretton GL; |
| Oy | 161 ATGATTTCGACGCCGGAAGGCTATGATCTTTCGAAGCAAGATGGCGGAATGCA 220 | DR | WRI; 2003-095291/82. | |
| Db | 61 AGCATTTTCGAACTGAGGTACATAGTTATGAAASCATCATGSGTCAGAAATGCA 120 | XX | P-PSDB; ADP05920. | |
| Oy | 221 CAGATCCCTCTGATAATGACATCACACTGTTGATCATGGATACTCGCGGGTAAG 280 | PT | New Proteus mirabilis polypeptides and polynucleotides, useful as reagents for diagnosis of bacterial disease, as components of anti-bacterial vaccines, as targets for antibiotic drugs, or as biocontrol agents for plants. | |
| Db | 121 CATATTCTCTGAAACAGCAATTAACTGGTGTATGGATACTTCAGGCAA 180 | PT | The invention relates to new Proteus mirabilis polypeptides and polynucleotides. The invention also relates to antibodies against the polypeptides, methods for producing the polypeptides, a method of generating vaccine for immunising an individual against <i>P. mirabilis</i> , a method for evaluating a compound for the ability to bind a <i>P. mirabilis</i> polypeptide and a method for screening test compounds for anti-bacterial activity. The polypeptides and polynucleotides are useful as molecular targets for diagnosing, preventing, and treating pathological conditions resulting from bacterial infection, as reagents for diagnosis of bacterial diseases, as components of antibiotic vaccines, as targets for antibacterial drugs or as bio-control agents for plants. This sequence represents a Proteus mirabilis polynucleotide of the invention. | |
| Oy | 281 AACGGTCCTCTGTTAGCGCGTGACTSGGGAGCAGCAAGAATGTTGCTTCG 340 | XX | Disclosure; SEQ ID NO 2033; 870pp; English. | |
| Db | 181 ATATGGCTGTACTTGCGCGCGACTGGAGCAAGAACATGTTGCTTCG 240 | CC | CC | |
| Oy | 341 ACTGGCCCTGACAACGAGTGTGATAAATTCTGGCCCTGAAATCGGCCCTGAAATGACTAC 400 | CC | CC | |
| Db | 241 ACTGGCCCTGATAACGAAAGTAGACAAATCTTCTGGCTTGGAAATCGGCCCTGAAATGATTAC 300 | CC | CC | |
| Oy | 401 ATCACCACAAACGGTCAACGGCGTGAATGACCTACTGTCCGT 460 | CC | CC | |
| Db | 301 ATCACCACAAACATTAAATCCGGCTGAATGACTATTCCGGCCGAACTCTTCGT 360 | CC | CC | |
| Oy | 461 ACCATGAACTCTGGTACTGTCAAGCAGAACTGGTACGGTGAAASCTAACATGTCAT 520 | CC | CC | |
| Db | 361 ACCATGAACTCTGGTAAATGTCAGCGAGAACGGCAAGCTGAGGATTAATTCAAC 420 | CC | CC | |
| Oy | 521 GGTGGGAACCTGACAAACGGCTGTTGATCGCCCTGAGGGAGCTACAG 580 | XX | Sequence 726 BP; 224 A; 137 C; 149 G; 216 T; 0 U; 0 Other; | |
| Db | 421 GGTGGGAATATGATATCATGAACTACGGCTCTCTGATAAGCCCGCTGGTGAGCGTAA 480 | XX | Query Match 48.6%; Score 450.8; DB 10; Length 726; Best Local Similarity 76.3%; Pred. No. 5e-139; Matches 554; Conservative 0; Mismatches 172; Indels 0; Gaps 0; | |
| Oy | 581 CTGGCGCGAGCGAGTTCGGCGCATGTTACTCTCTGAAACCCGGAAATTCAG 640 | DB | 92 GTGACAAACATCAGACGCCGCACATTCTCTGTTAGTGTGAGAGCTGGTACACGGAC 151 | |
| Db | 481 TTGCCACCGAGCGAGTTCGGCGCATGTTACTCTCTGAAACCCGGAAATTCAG 540 | DB | 1 GTGACAAACATCAGACGCCGCACATTCTCTGTTAGTGTGAGAGCTGGTACACGGAC 60 | |
| Oy | 641 TCCGTGCTGAACTGTCTGAAAGATGACCGCGCTGAGTAAACGGCACACCGTACT 700 | QY | 152 AGTTGAAAGATTTCGAGGGGCTATGAGTTTCTGAAAGCAGATGGCG 211 | |
| Db | 541 ACACGGCAGATTACTCAAAGATGACAGGCTGTGATGAACTCAACGGCTACT 600 | DB | 61 ACCTGAAAGCATATTGAGAGCTGGTATATCTACAGGACACTGATGGCAAC 120 | |
| Oy | 701 GTGACCGTGAATCCCGCTATCTGAAACATTGATCTACGGCGATACGGCGAA 760 | QY | 212 GAATGATCAGATCTCTGAAATGACATCAACGGCATCTGGATATCACTG 271 | |
| Db | 601 GTGATGTAACGATTCCTGTAAACATTGATCTGAGCTGGCAATGAGTGGAA 660 | DB | 121 GAGATGATAATATTCTCTGCTCCACCATGATATCACTGTCGTCATTAGTATACTT 180 | |
| Oy | 761 ATCATGCCACCATTCAGGGAGGTATCGTTCTGGTGTATCTGAAAGATAA 817 | QY | 272 CCCGTTAGAACGGCTCTCTGTTAGCGCGTGAATCTGGCGAGCAGGAATGTTGCTTA 331 | |
| Db | 661 ATTATGCCACCATTCAGGGAGGTATCGTTCTGGTGTATCTGAAAGATAA 717 | DB | 181 CCTGGTAAATGGTCTCTTATAGCCGTGATAGTGAACAGTAAAGTGTGCTTA 240 | |
| RESULT 8 | | QY | 332 ATGTTCTCTGACTGGCGCTGACAGAGTGTGATAATTCGGCGTGAATATCGCTGAACATGGCTGA 391 | |
| ID ADP01748 | | DB | 241 ATGTTCTCTAAGCTGGCTGTGATGAGGTTGATAATCTTGGGCC 300 | |
| AC ADP01748; | | QY | 392 GATGACTACATCACCACCGTCACCCGCGTGAATGAGTGGCAGCACTA 451 | |
| XX 12-FEB-2004 (first entry) | | DB | 301 GATGATTAACATCAACATTAACTTAACTCTGTAATTAACCTCGTGTGCTTA 360 | |
| DE Bacterial polynucleotide #2033. | | QY | 452 CTGCGCGTACATGAACTGGTACTGTCAAGGAGAAGCTGGTGTAGCTGAAGCTAC 511 | |

| | | | | | |
|--|---|--------|---------|----|--|
| WP | AAT>2063_15 | 150001 | 161000 | KW | antibacterial; drug design. |
| WP | AAT>2063_16 | 160001 | 171000 | XX | |
| WP | AAT>2063_17 | 170001 | 181000 | OS | Haemophilus influenzae. |
| WP | AAT>2063_18 | 180001 | 1830121 | XX | |
| Query Match | 34.5%; Score 319.4; DB 2; length 110000; | | | PN | |
| Best Local Similarity | 65.8%; Pred. No. 5.4e-94; Mismatches 246; Indels 3; Gaps 1; | | | XX | |
| Matches 480; Conservative 0; Mismatches 246; Indels 3; Gaps 1; | | | | PD | 27-SEP-2001. |
| Qy | 85 AATTAGCTGACAACTGAGACGCCACATCTTATCGTTGAAGCAGAGTGGPAC | 144 | | XX | 21-MAR-2001; 2001WO-US009180. |
| Db | 36640 AGTATAAGGATGCAAATGACTCCAAAATCTCGTTGTGAGAATGTCAC | 36581 | | PR | 21-MAR-2000; 2000US-0191078P. |
| Qy | 145 ACCAACACAGTGAAGTATTTCGAGGGAGCTATGTTGAAAGCAGAGTGGPAC | 144 | | PR | 23-MAY-2000; 2000US-020778P. |
| Db | 36580 TCGAAATACGCTTAAGGGATTCTCGTTGAGGGAGGATATGATGTTGAGCAGAAA | 36521 | | PR | 23-OCT-2000; 2000US-024255P. |
| Qy | 205 TGCAGCGAAATCATCGATCTCTCTGATATGACATCAACCTGGTATGATGAT | 264 | | PR | 22-DEC-2000; 2000US-025365P. |
| Db | 36520 TGGTGTGAAATGATCATATATGGCAATCATATATTATGGTGTGATGGAT | 36461 | | PR | 16-FEB-2001; 2001US-0257938P. |
| Qy | 265 CAATCTCCGGTAAAGAACGGRCTCTGTTAGSGGTTAACTGGCAATGCT | 324 | | XX | |
| Db | 36460 TAATTACAGGCAAAACGGCTTATTTGGCAGAGAACCTCGTGTGAGAATAGCT | 36401 | | PA | (ELIT-) ELITRA PHARM INC. |
| Qy | 325 TGGTGTGATGTTCTGACTGGCGTGACAAAGACTGATTAATTCTGGCTGAAT | 384 | | XX | |
| Db | 36400 ACCTCTTATTGTTTAACTGTTGAGATTAATGAGTTAGTTAGTTAGTTGGAAT | 36341 | | PI | Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ; |
| Qy | 385 CGGTGAGATGACTACATCACCAAAACCTTCAACCGGGTGAATGAGGATCTGGCACG | 444 | | XX | Yamamoto RT, Xu HH; |
| Db | 36340 TGGTGGCGATGACTATTGACTAAGCTTAACTCTAGAAGACTGACTATCGCGCAG | 36281 | | DR | WPI; 2001-611495/70. |
| Qy | 445 CAACTCTGTCGGTACATGATCTGGTACTGTCGCGGAAACTGTTGAGCTGA | 504 | | XX | P-PSDB; AAU35550. |
| Db | 36280 TAATCTATGTCATGTTGAAATG--CCCATGAGAAATAGAAATACATTGTCGAGA | 36224 | | PS | |
| Qy | 505 AACCTACAGTCAATGTTGGAACTGCACTCACAGCGTCGTTGATCGCCCTGA | 564 | | XX | |
| Db | 36223 ATTCATATGTTTAATGGTTGAAATTGACCTAAATGACAGCTTAATGACACAGA | 36164 | | PT | New polynucleotides for the identification and development of antibiotics, comprise sequences of antisense nucleic acids. |
| Qy | 565 TGCGGAGGAGTACAAGCTGCCGGCTGAGTCGGTCCGGCCATCTTCACTCTGTGAAA | 624 | | XX | Claim 27; SEQ ID NO 7046; 51PP; English. |
| Db | 36163 AGGACAGAGATCAACCTCTCGCAGTGATTGTCGCAATGTTCTGTGAAA | 36104 | | CC | The invention relates to antisense inhibitors of genes essential to prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes themselves and the encoded proteins. The prokaryotes used are Escherichia coli, <i>Staphylococcus aureus</i> , <i>Salmonella typhi</i> , <i>Klebsiella pneumoniae</i> , <i>Pseudomonas aeruginosa</i> and <i>Enterococcus faecalis</i> . The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence encodes an essential prokaryotic cellular proliferation protein. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences |
| Qy | 625 CCCAGGAAATTCAGTCGCCGCTGACTGCTGAGAAGAACCGGCCGAGCGAA | 684 | | CC | |
| Db | 36103 TCCAGGAAATTCAGCAAGGCCGAGAATGCTCCGACGAGTAA | 36044 | | CC | |
| Qy | 685 ACCGCACCGACCGTACTGTAGACCTGCCGTTCTGAACATTCTGCAATCTAC | 744 | | CC | |
| Db | 36043 ACCTCAGGATCTGTCACCGTAGCTCACATTCACAGTATGAGAAACATTGAGATCA | 35984 | | SQ | Sequence 711 BP; 246 A; 113 C; 209 T; 0 U; 0 Other; |
| Qy | 745 GCGGATAAGCCGAAATCATGCCACATTACGGTCAAGGTTATGCTCTGCGGTA | 804 | | SO | Query Match 34.1%; Score 316.2%; DB 4; Length 711; Best Local Similarity 66.2%; Pred. No. 3.6e-4; Mismatches 238; Indels 3; Gaps 1; |
| Db | 35983 TCCAAATCTCCAATATCATATGACTATACATGGAGAGGCTATGTTGAGA | 35924 | | Qy | 1.01 ATCCAGGCCACATCTCTCGTTGAGAGCGAGTGGTACACCCACACGTTGAA 160 |
| Qy | 805 TCTGGAAGA | 813 | | Db | 1 ATGACTACTCCAAATCTCGTTGAGATGAAATGTCACCTCGAAATAGCTAA 60 |
| Db | 35923 TATTGAAATA | 35915 | | Qy | 1.61 ATGATTTCGAGCGGGAGGCTATGAGTTTCTGAGAGCAGATGCGGAAATGCT 220 |
| RESULT 11 | | | | Db | 61 GGATTTGGCAATCATATAATTATTTGGTTGATGAAATGCTGAAATGCT 120 |
| AAS5409 | | | | Qy | 221 CAGATCTCTCTGACATGACATCACCTGGTATGATGTTGAGAATGCTGTTAG 280 |
| ID AAS5409 standard; DNA; 711 BP. | | | | Db | 121 CATATATGGCAATCATATAATTATTTGGTTGATGAAATGCTGAAATGCT 180 |
| XX AAS5409; | | | | Qy | 281 AACGGTCTCTGTTAGGGTCAACTCGCGAGCAGCGGAATGTTGCTGCTG 340 |
| DT 13-FEB-2002 (first entry) | | | | Db | 181 AACGGTATATTGGAGAGAACCTCGTGTGAGAAATGCTTACCTTATTGTTA 240 |
| DE Haemophilus influenzae DNA for cellular proliferation protein #191. | | | | Qy | 341 ACTGGCCGTGACAACGAGTCGATAAAATTCTGGCTCGAAATCGGTGAGATGACTAC 400 |
| KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic; | | | | | |

Db 241 ACTGGTGGAGATACTGAGTGGATAAATTAGTTAGTTAGGAAATTGGGGCGGAGTACT 300
 QY 401 ATCACCACACCGTCACCGCGCTGAACCTGAGATTCGTGCAAGCAACCTACTGTCCCGT 460
 Db 301 TTGACTTAAGCTTAACTGACCTAGAATCTGAGTACTATCCGGCAAGTATCTTGATGT 360
 QY 461 ACCATGAAATCTGGTACTTCAGGAGAACGAGCTGAGTGTAGCGTGAAGCTACAGTCAAT 520
 Db 361 GCAATG--CCGCATCACAGAAAAGAAAATACATTGGTGGAGAACCTACAGTCAAT 417
 QY 521 GGTTGGAGACTGGACATCACAGCGGTTGATCGGCCCTGATSGCGACAGTCAAG 580
 Db 418 GGTGGAATATGACTAAATAGCCACAGTTAACTACCGAGAACGAAATTCGAA 477
 QY 581 CTGCGCGGAGCGAGTCGGCGCATGCTCTACTCTGAGAACCCAGGAAATTGCG 640
 Db 478 CTTCTCTCGAGTCAGTGTGCAATTTGCAATTTGTTGAGAACATCCAGAAATTGCAA 537
 QY 641 TCCCGTGTGAACCTGAGAATACCGCCGTTGAGCTGAACCGCAACGGTACT 700
 Db 538 ACCCGCGAGAAATGCTGAAAAAATGACCGACGAGAGTAAACCTGAGGATGTTAC 597
 QY 701 GTAGACCGTGGACGATCCCGGTTGATTAACATTGATCTACCGGGTACCCGGA 760
 Db 598 GTAGATCTCACATTCTGAGTACAGAAACATTGAGATCATCCAACTCCAAAT 657
 QY 761 ATCATCCCCACATTCTGAGTACAGAAACATTGAGATCATCCAACTCCAAAT 813
 Db 658 ATCATATGACTATACATGAGAAGGGTATCTGTTTGGAGGATATGATA 710

RESULT 12

ACA34287
 ID ACA34287 standard; DNA; 711 BP.
 XX AC ACA34287;
 XX DT 19-JUN-2003 (first entry)
 XX DR Prokaryotic essential gene #115944.
 KW Antisense; ds; prokaryotic essential gene; cell proliferation; drug design; gene.
 XX OS Haemophilus influenzae.
 PN WO200277183-A2.
 PD 03-OCT-2002.
 XX P1 21-MAR-2002; 2002WO-US009107.
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-0094893.
 PR 25-OCT-2001; 2001US-034293P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 PA (ELIT-) ELITRA PHARM INC.
 XX P1 Wang L, Zamudio C, Malone C, Haselbeck R, Olsen KU, Zyskind JW;
 Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 DR WPI; 2003-029926/02.
 DR P-PSDB; ABU30417.
 PT New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.
 PS Claim 14; SEQ ID NO 22157; 176pp; English.

Sequence 711 BP; 246 A; 113 C; 143 G; 209 T; 0 U; 0 Other;
 SQ Sequence 711 BP; 246 A; 113 C; 143 G; 209 T; 0 U; 0 Other;
 Query Match 34.1%; Score 316.2%; DB 8; Length 711;
 Best Local Similarity 66.2%; Pred No. 3; 6e-99;
 Matches 472; Conservative 0; Mismatches 238; Indels 3; Gaps 1;
 QY 101 ATCGAGACCCGGACATCTTCTTCTGAGACGACTTGGTACACGGACACGGTGA 160
 Db 1 ATGACTACTCCAAAATCTCGTGTGTTGAGATGAAATGTCACTCGAAATACGTTAA 60
 QY 161 AGTATTTGAGGGATGATGTTCTGAGACGACTTGGTACACGGACACGGTGA 220
 Db 61 GGGATTTGAGCGGAGGATGATGTTGAGGAGAAATGGTGTGAAATGCA 120
 QY 221 CAGATCCCTCTCATATGACATCAACCTGGTACATGATCATGATATCAATGCGGGTAG 280
 Db 121 CATATATGGCAATCATATAATTATTTGGTGTGATGGATGATTAATTACAGGAA 180
 QY 281 AACGGTTCTTCTTGGCTGAGCTGCGCGAGGCGCATGTTGCGTGTGATGTTCTG 340
 Db 181 AACGGCTTATATGGCAAGAAGAACCTCGTGGAGATTAAGCTTACCTTATTTTA 240
 QY 341 ACTGGCGGAGAACGAGTGGATTAATCTCGCTGGAAATGGTGTGAGTACTAC 400
 Db 241 ACTGGTGGAGATACTGAGTAAATTAGTTAGTTGGAGAACATGGTGTGAGTACTAC 300
 QY 401 ATCACCAACCGTTCACCGCGGAACGGTACTGAGATTCGGCTGGAACTACTGTCCCGT 460
 Db 301 TTGACTTAAGCTTAACTGACCTAGAATCTGAGTACTATCCGGCAAGTATCTTGATGT 360
 QY 461 ACCATGAAATCTGGTACTTCAGGAGAACGAGCTGAGTGTAGCGTGAAGCTACAGTCAAT 520
 Db 361 GCAATG--CCGCATCACAGAAAAGAAAATACATTGGTGGAGAACCCAGGAAATTGCG 417
 QY 521 GGTTGGAGACTGGACATCACAGCGGTTGATCGGCCCTGATGGCAGTAATCTTGATGT 580
 Db 418 GGTGGAATATGACTAAATAGCCACAGTTAACTACCGAGAACGAAATTCGAA 477
 QY 581 CTGCGCGGAGCGAGTCGGCGCATCTCTACTCTGAGAACCCAGGAAATTGCG 640
 Db 478 CTTCTCTCGAGTCAGTGTGCAATTTGCAATTTGTTGAGAACATCCAGAAATTGCAA 537

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation; (7) identifying the activity of a gene in an operon required for the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing a compound that influences the activity of the product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this target did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

QY 641 TCCCGTGTGACTCTGAGAATGACGGCGTGAACCGCACCGTACT 700
 XX Sequence 1110 BP; 256 A; 291 C; 296 G; 267 T; 0 U; 0 Other;
 Db 538 ACCGGCGAGGAATGCTGAAAMATGACGGAGAGTGTAACCTCGGGATGTTAC 597
 QY 701 GTAGAGCGTGAAGATCGCGCGTATTGTTAACATTCGAACTACGCCGATAGCCGAA 760
 Db 598 GTAGATGTCACAATTCGAGGTATCGAGATCATCCATACTCCAAT 657
 QY 761 ATCATCGCCACCAATTCAGGTGAGGTATCGCTCAGGTATCGAGA 813
 Db 658 ATCATATGACTATGAGAGGTATCGTTTGGAGATATGATA 710
 RESULT 13
 AAS82422 ID AAS82422 standard; cDNA; 1110 BP.
 AC XX AAS82422;
 DT 13-FEB-2002 (first entry)
 DE DNA encoding novel human diagnostic protein #18226.
 KW human; chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX Homo sapiens.
 OS Homo sapiens.
 PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US008631.
 XX PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 PA (HYSEQ-) HYSEQ INC.
 PT Drmanac RT, Liu C, Tang YT;
 XX PI XX
 DR XX
 DR XX
 PT New isolated Polynucleotide and encoded polypeptides, useful in
 diagnostics, forensics, gene mapping, identification of mutations
 PT diagnostic for genetic disorders or other traits and to assess
 PT biodiversity.
 XX PS Claim 1; SEQ ID NO 18226; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probe, polymerase chain
 reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 diagnostics, forensics, gene mapping, identification of mutations
 responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other type of data and products dependent on DNA and
 CC amino acid sequences. AAS61917-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

QY 544 CCGTCTGTTGATCGCCCTGATGGGGAGTA-CAACTGCAGGAGTCGG 602
 XX Query Match 25.1%; Score 232.8; DB 5; Length 1110;
 Best Local Similarity 98.8%; Pred. No. 2.9e-66; 0; Mismatches
 Matches 245; Conservative 0; Indels 1; Gaps 1;
 Db 761 CCGTCTGTTGATCGCCCTGATGGGGAGTCGG 820
 QY 603 CCATGCTTCACTTGTTGAAACCCAGCAAAATTCTGCTCCCTGCTGAACTGTGAGA 662
 Db 821 CCATGCTTCACTTGTTGAAACCCAGCAAAATTCTGCTCCCTGCTGAACTGTGAGA 880
 QY 663 AATGAGCCGCGCTGAGATCACGCCGAGACGGTAGCTGAGTCAGTCGGCTA 722
 Db 881 AATGAGCCGCGCTGAGATCACGCCGAGACGGTAGCTGAGTCAGTCGGCTA 940
 QY 723 TTGCTAACATTGAGATCTACGCCGAAATCATGCCACCATTCACCGTG 782
 Db 941 TTGCTAACATTGAGATCTACGCCGAAATCATGCCACCATTCACCGTG 1000
 QY 783 AACGTTAT 790
 Db 1001 AAGTTCAT 1008

RESULT 14
 ABQ90117 ID ABQ90117 standard; DNA; 729 BP.
 XX ABQ90117;
 XX PR 01-OCT-2002 (first entry)
 DR XX M. capsulatus gene #102 for DNA array.
 XX Micro array; gene; ds; differential expression; gene expression.
 XX OS Methylococcus capsulatus.
 PN XX WO200255655-A2.
 PD XX 18-JUL-2002.
 PF XX 14-JAN-2002; 2002WO-N0000019.
 PR XX 12-JAN-2001; 2001NO-00000235.
 PR XX 12-JAN-2001; 2001NO-00000239.
 PA (UNIF-) UNIFOB STIFTESN UNIV BERGEN.
 PA XX Birkeland NK, Edhammar I, Jonassen I, Jenden HB, Lien T,
 PR Lillehaug JR, Lüssius I, Eisen JA, Fraser CM, Durkin AS;
 PR Salzberg SL;
 DR XX WPI; 2002-557818/59.
 PS XX Novel DNA array useful for determining differential expression of
 PT Methylococcus capsulatus genes, comprises polynucleotides or
 PT oligonucleotides representative for a selective number of Methylococcus
 PT capsulatus genes.
 PS XX Claim 19; Page 99; 678pp; English.

The invention relates to a novel DNA array giving a representation of a
 CC number of Methylococcus capsulatus genes. The method of the invention is
 CC useful for determination of the differential expression of the genes of
 M. capsulatus, and for studying gene expression on a genomic scale and in
 CC gene expression assays of M. capsulatus genes. The sequences shown in
 CC ABQ90116-ABQ91855 represent M. capsulatus genes for use in arrays of the

Search completed: January 22, 2005, 12:42:09
Job time : 556 secs

This Page Blank (uspto)

Copyright (c) 1993 - 2005 Compugen Ltd.
GenCore version 5.1.6

OM nucleic - nucleic search, using SW model
Run on: January 22, 2005, 12:15:23 ; Search time 105 Seconds
(without alignments)

6275.250 Million cell updates/sec

Title: US-10-613-990A-31
Perfect score: 927
Sequence: 1 gtcatgttacggcgatcatg.....gttaactttttactggctt 927

Scoring table: IDENTITY_NUC Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:
1: /cgn2_6/ptodata/lina/5A_COMB.seq: *
2: /cgn2_6/ptodata/lina/5B_COMB.seq: *
3: /cgn2_6/ptodata/lina/6A_COMB.seq: *
4: /cgn2_6/ptodata/lina/6B_COMB.seq: *
5: /cgn2_6/ptodata/lina/PCUS_COMB.seq: *
6: /cgn2_6/ptodata/lina/backtiles.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | DB ID | Description |
|------------|-------|--------------------|-------|---|
| 1 | 577.2 | 62.3 | 726 | 4 US-09-489-039A-1812 Sequence 1812, Ap |
| 2 | 450.8 | 48.6 | 726 | 4 US-09-543-681A-2033 Sequence 2033, Ap |
| c | 3 | 319.4 | 34.5 | 1830121 4 US-09-557-884-1 Sequence 1, Appl |
| c | 4 | 319.4 | 34.5 | 1830121 4 US-09-643-990A-1 Sequence 1, Appl |
| c | 5 | 126.6 | 13.6 | 837 4 US-09-232-991A-16417 Sequence 16417, A |
| c | 6 | 126.6 | 13.6 | 2420 4 US-09-232-991A-16417 Sequence 16417, A |
| c | 7 | 126.6 | 13.6 | 2420 4 US-09-232-991A-16417 Sequence 16417, A |
| c | 8 | 126.6 | 13.6 | 3675 4 US-09-232-991A-16309 Sequence 16309, A |
| c | 9 | 120.6 | 13.0 | 990 4 US-09-232-991A-1517 Sequence 11517, A |
| c | 10 | 120.6 | 13.0 | 2166 4 US-09-252-991A-1172 Sequence 1172, A |
| c | 11 | 113.8 | 12.3 | 2004 4 US-09-232-991A-11865 Sequence 11865, A |
| c | 12 | 107.4 | 11.6 | 1018 4 US-09-418-880-1 Sequence 1, Appl |
| c | 13 | 98.6 | 10.6 | 729 4 US-09-634-238-133 Sequence 133, Appl |
| c | 14 | 97.2 | 10.5 | 762 4 US-09-232-991A-16307 Sequence 16309, A |
| c | 15 | 96.2 | 10.4 | 705 3 US-09-14-01C-226 Sequence 226, App |
| c | 16 | 94.6 | 10.2 | 702 4 US-09-710-279-479 Sequence 479, App |
| c | 17 | 94.6 | 10.2 | 3389 4 US-09-710-279-3914 Sequence 3914, App |
| c | 18 | 92.6 | 10.0 | 792 4 US-09-252-991A-10543 Sequence 10543, A |
| c | 19 | 92.6 | 10.0 | 804 4 US-09-252-991A-10323 Sequence 10323, A |
| c | 20 | 91.4 | 9.9 | 702 4 US-09-082-077-2 Sequence 2, Appl |
| c | 21 | 91.4 | 9.9 | 3731 4 US-09-082-077-1 Sequence 1, Appl |
| c | 22 | 91.4 | 9.9 | 9456-171B-87 Sequence 87, Appl |
| c | 23 | 91.4 | 9.9 | 9425 4 US-08-781-986A-87 Sequence 87, Appl |
| c | 24 | 87.6 | 9.4 | 783 4 US-09-483-039A-372 Sequence 372, Appl |
| c | 25 | 87.4 | 9.4 | 945 4 US-09-232-991A-3092 Sequence 3092, Appl |
| c | 26 | 87.4 | 9.4 | 978 4 US-09-252-991A-2878 Sequence 2878, Appl |
| c | 27 | 87.4 | 9.4 | US-09-252-991A-2990 Sequence 2990, Appl |

RESULT 1
US-09-489-039A-1812
; Sequence 1812, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITL OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709 . 200401
; CURRENT APPLICATION NUMBER: US/09/489, 039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117, 747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1812
; LENGTH: 726
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae

ALIGNMENTS

Query Match 62.3%; Score 577.2; DB 4; Length 726;
Best Local Similarity 87.2%; Pred. No. 1.le-197; Mismatches 93; Indels 0; Gaps 0;
Matches 633; Conservative 0; Sequence 1812, Ap

QY 92 GTAGCAACATGAGACCCGCACATTCTATCGTTAGAGCAGGTGGTAAACCGAAC 151
1 GTAGCAACATGAGACCCGCACATTCTATCGTTAGAGCAGGTGGTAAACCGAAC 60

QY 152 ACCTTGAAAGATTTCGAAGCGGAGGCTTGATGTTTGAAGGAGACATGSGCG 211

QY 61 AGCTTAAGAGTTTCTGAAGGAGTGTAGATGATTCGAAGGACCTGGCG 120

QY 212 GAATGATCAGATCCTCTGATCATCACATGATCATGGATCACTCTG 271

Db 121 GAAATGATCAGATCCTCTGATCATGGATCACTCTG 180

QY 272 CCGGGTAGAACCGCTCTCTGTTAGCGCGAGGGAGATGTGCGGTG 331

Db 181 CCGGGTAGAACCGCTCTCTGTTAGCGCGAGGGAGATGTGCGGTG 240

QY 332 ATCTTCGACTGGCCGAGCAAGGAAGTCTGATAATTCGGCCCTGAAATCGGCGCA 391

Db 241 ATCTTCGACTGGCCGAGCAAGGAAGTCTGATAATTCGGCCCTGAAATCGGCG 300

QY 392 GATGACACATCACCACCGCTCAACCGCGCTGACCTGGCGAGGTGACCTG 451

Db 301 GAGCACTATACCTAACCGTGTACCCGGCGTGAACCTGACTATCCCGCGCGCACTG 360

QY 452 CTGGCCCTACCTGATGAGCTGGTACTGTCAGGGAGAACGTTGAGCTGTTAGGGTGAAGGTAC 511

RESULT 2
US-09-543-681A-2033
; Sequence 2033, Application US/09543681A
; Patent No. 6505709
; APPLICANT: GARI BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709_1002-001
; CURRENT APPLICATION NUMBER: US 09/543, 681A
; PRIORITY FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIORITY FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO: 2033
; LENGTH: 726
; TYPE: DNA
; ORGANISM: Proteus mirabilis
; US-09-543-681A-2033

Query Match 49.6%; Score 450.8; DB 4; length 726;
Best Local Similarity 76.3%; Pred. No. 5e-152;
Matches 554; Conservative 0; Mismatches 172; Indels 0; GapB 0;

QY 92 GTAGGAAACATGGAGACCCCGCACATCTTATCGTTGAGACGAGTGTGTAACACGGAC 151
Db 1 GTAGCAAAATGCGAACCCCGCACATCTGATGTTGAGATGAGTAGTACTCGTAAT 60
QY 152 ACGTTGAAAGTATTGAGACGGGAAGCTGTGAGCATGTTGAGACAGATGGGGCG 211
Db 61 ACCCTGAAAGCTATTGCGAAGCTGTGAGGGTATTCGACAGGAAGCCACTGATGCCAC 120
QY 212 GAAATGCTCGATCTCTCTGATATGACATCACCTGGTACATCATGGATACTCACTG 271
Db 121 GAGATGCTAAATATTCTGTCGACCGATGATCATGGTATATGGATAATTAACTT 180
QY 272 CCGGGTAAAGAACGGTCTCTGTAGCTGGTGAACCTGGCAGCAGGGAAAGTTGGTG 331
Db 181 CCCTGTTAAATGGTCTCTTATAGCCGTGTTAGTAGGACAGGAAGTGTGCAATA 240
QY 332 ATGTTCCCTGACTGGCCCTGACAACGAGTCGATAAATTCTGGCTCGAATCGGCA 391
Db 241 ATGTTCCCTAACAGGTCTGTGATATGAGTTGATAAACTTAGGCTTGAAATTGGGC 300
QY 392 GATGACATCACCAACCGTCACCCGGTGAACTGAGATGTTGCAACGCCACCTA 451

RESULT 2
US-09-543-681A-2033
; Sequence 2033, Application US/09543681A
; Patent No. 6505709
; APPLICANT: GARI BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709_1002-001
; CURRENT APPLICATION NUMBER: US 09/543, 681A
; PRIORITY FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIORITY FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO: 2033
; LENGTH: 726
; TYPE: DNA
; ORGANISM: Proteus mirabilis
; US-09-543-681A-2033

Query Match 49.6%; Score 450.8; DB 4; length 726;
Best Local Similarity 76.3%; Pred. No. 5e-152;
Matches 554; Conservative 0; Mismatches 172; Indels 0; GapB 0;

QY 92 GTAGGAAACATGGAGACCCCGCACATCTTATCGTTGAGACGAGTGTGTAACACGGAC 151
Db 1 GTAGCAAAATGCGAACCCCGCACATCTGATGTTGAGATGAGTAGTACTCGTAAT 60
QY 152 ACGTTGAAAGTATTGAGACGGGAAGCTGTGAGCATGTTGAGACAGATGGGGCG 211
Db 61 ACCCTGAAAGCTATTGCGAAGCTGTGAGGGTATTCGACAGGAAGCCACTGATGCCAC 120
QY 212 GAAATGCTCGATCTCTCTGATATGACATCACCTGGTACATCATGGATACTCACTG 271
Db 121 GAGATGCTAAATATTCTGTCGACCGATGATCATGGTATATGGATAATTAACTT 180
QY 272 CCGGGTAAAGAACGGTCTCTGTAGCTGGTGAACCTGGCAGCAGGGAAAGTTGGTG 331
Db 181 CCCTGTTAAATGGTCTCTTATAGCCGTGTTAGTAGGACAGGAAGTGTGCAATA 240
QY 332 ATGTTCCCTGACTGGCCCTGACAACGAGTCGATAAATTCTGGCTCGAATCGGCA 391
Db 241 ATGTTCCCTAACAGGTCTGTGATATGAGTTGATAAACTTAGGCTTGAAATTGGGC 300
QY 392 GATGACATCACCAACCGTCACCCGGTGAACTGAGATGTTGCAACGCCACCTA 451

RESULT 3
US-09-557-884-1/C
; Sequence 1, Application US/09557884
; Patent No. 650581
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Mark
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB186P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 183011 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

US-09-557-884-1
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Query Match 34.5%; Score 319.4; DB 4; Length 1830121;
Best Local Similarity 65.8%; Pred. No. 4.4e-102;
Matches 480; Conservative 0; Mismatches 246; Indels 3; Gaps 1;

OY 85 AATTAGGTGCAACATCGAGACCCGCACATCTATCGTGAAGACGAGTTGGAAC 144
Db 936640 AGTATAAGTAGTGAAGAATGACTACTCCAAAATCTCGTTGTGAAGATTCAC 936581
OY 145 ACCGACAGCTTGAAAGATTTTCGAAGGGAGGCTATGATTTTCGAAACGCCAGA 204
Db 936580 TCGAAATACGTTAACGGATTITGAAAGGGATATGATGTTGAAAGCAGAA 936521
OY 205 TGGCGCGAAATCATCGATCCCTCTGAAATAGACATCAACCTGGATCATGGAT 264
Db 936520 TGGTGTGAAATGATCATATATGGCAATCATATAATTAATTTGGTTGAT 936461
OY 265 CAACTGCGGGTAAAGACGGCTCTCGTAGCCGTGAACTGGGGAGGGAAAT 324
Db 936460 TAATTACCAAGCRAAACCGCTTATTATGGAAAGACATCGTGAAGAATTA 936401
OY 325 TGGTTGATGTTCTGACTCTGCGCTGACAAACGAGACTGCGTGAATAAAATTCTGGCTCGAAAT 384
Db 936400 ACCCTCTTAAATTAACTGTGTGAGATATGAGTTGAAATTAGTTGGAAAT 936341
OY 385 CGGTGAGATGACTACATACCAACCGTTAACCCGGTGAACTGCGATTCGGCAG 444
Db 936340 TGGTGCCTGAAATGACTATTGACTAAGCCTTTAACCTTGAAGACTGACTATCGCGCAG 936281
OY 445 CAACCTACTGCGTACATGAACTGGTACTGTCAAGGAACTGCGTGAACGGTGA 504
Db 936280 TAATCTATGCACTGCTGCAATG--CCCATCAAGAAAGAAATACTTGTGAGA 936224
OY 505 AANGTACAGTTCATGGTGGAACTGCGACATCAACAGCCGTTGCGTGTGATCGGCCCTGA 564
Db 936223 ATTCATGCTTAACTGGAAATTAGACCTAAACACAGTTAACACCGA 936164
OY 565 TSGCAGACGATGAGCTGGCGGAGGAGTCGGCAGGAGTCGGTCAAGGAAAT 624
Db 936163 AGGACAAGAGATCAACTCTCGCAGTCA 936104
OY 625 CCCAGGCAAAATTCACTGGCTGACTCTGCTGAGAAATGACCGCCGAGCTCTGTGAA 684
Db 936103 TCCAGGAAATTCGAAAGGCCGAGAAATGCTGAAAAATAACCGSAGAGTAA 936044
OY 685 ACCGACAGCAGCTACTGTGAGCGTACCGTCCCGCTGTTCTGAAACATTCTGATCTAC 744
Db 936043 ACCTCAGGATGCTACCGTACAGTCACAAATTGCAAGTATCAGAAACCTTGGAGATCA 935984
OY 745 GCGGAGATGCCGAAATCATCGCACCATCAGGTGAGGGTTACCTCTGGGTTGA 804
Db 935983 TCCCAATCTCCAAATATCATATTGACTATCAGTGGAGGGCTATCGTTTGTGAGA 935924
OY 805 TCTCGAAGA 813
Db 935923 TATTGAAATA 935915

NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville,
STATE: MD
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643,990A
FILING DATE: 23-Aug-2000

CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07

APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21

ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover

REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB186P1C1

TELECOMMUNICATION INFORMATION:
TELEPHONE: 310-309-5790
TELEFAX: 310-309-8439

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDBEADNESS: double

TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-643-990A-1
Query Match 34.5%; Score 319.4; DB 4; Length 1830121;

Best Local Similarity 65.8%; Pred. No. 4.4e-102;
Matches 480; Conservative 0; Mismatches 246; Indels 3; Gaps 1;

OY 85 AATTAGGTGCAACATCGAGACCCGCACATCTATCGTGAAGACGAGTTGGAAC 144
Db 936640 AGTATAAGTAGTGAAGAATGACTACTCCAAAATCTCGTTGTGAAGATTCAC 936581
OY 145 ACCGACAGCTTGAAAGATTTTCGAAGGGAGGCTATGATTTTCGAACTCGA 204
Db 936580 TCGAAATACGTTAACGGATTITGAGCGGAGGATATGATGTTGAAAGCAGAA 936521
OY 205 TGGCGCGAAATCATCGATCGAGATCTCTGAAATAGACATCAACCTGGATCATGGAT 264
Db 936520 TGGTGTGAAATGATCATATATGGCAATCTAAATTAAATTGGTTGTGATGGAT 936461
OY 265 CAACTCTGGGTTAGGACGGTTCTCTGGCTGACACAGGTGATAAAATTCTGGCTCGAAAT 384
Db 936460 TAATTACCGAAACAGGGCTTATATGGCAAGAGACCTCGTGAAGAATTA 936401
OY 325 TGGTTGATGTTCTGCTGGCGTGAACACAGGTGATAAAATTCTGGCTCGAAAT 384
Db 936400 ACCCTCTTAACTGGTGAAGATAATGAGTGTGATAAAATTGTGTTGGAAT 936341
OY 385 CGGTGAGATGACTACATCACCACCGCTAACCCGGTGAACTGAGATTGTGAGCG 444
Db 936340 TGGTGCCTGAAATGACTATTGACTAAGCCTTTAACCTGGAGACTGACTATCGCGCAG 936281
OY 445 CAACCTACTGCGTACCGTACCGTCAACCGCTAACCCGGTGAACTGAGATTGTGAGCG 504
Db 936280 TAATCTATGCACTGCTGCAATG--CCCATCAAGAAAGAAATACTTGTGAGA 936224
OY 505 AACCTACAGTCAATGCTGGAACTGCGACATCAACAGCCGTTCTGTGATGGCCCTGA 564
Db 936223 ATTCTATGCTTAACTGGAAATTAGACCTAAACACAGTTAACACCGA 936164

CURRENT FILING DATE: 1999-02-18
 PRIORITY APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIORITY APPLICATION NUMBER: US 60/094,190
 PRIORITY FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO: 16417
 LENGTH: 837
 TYPE: DNA
 ORGANISM: *Pseudomonas aeruginosa*
 SEQ ID NO: 09-252-991A-16417
 LENGTH: 2430
 SEQ ID NO: 09-252-991A-16171
 LENGTH: 2430
 TYPE: DNA
 ORGANISM: *Pseudomonas aeruginosa*
 Best Local Similarity 52.9%; Pred. No. 2.4e-34;
 Matches 295; Conservative 0; Mismatches 260; Indels 3; Gaps 1;
 Query Match 13.6%; Score 126; DB 4; Length 837;
 Best Local Similarity 52.9%; Pred. No. 1.1e-34;
 Matches 295; Conservative 0; Mismatches 260; Indels 3; Gaps 1;
 Query Match 13.6%; Score 126; DB 4; Length 2430;
 Best Local Similarity 52.9%; Pred. No. 2.4e-34;
 Matches 295; Conservative 0; Mismatches 260; Indels 3; Gaps 1;
 Qy 243 TCAACCTGGTATCATGATAATCACTGCGGGTAGAACGGCTCTGTAGCGGTG 302
 Db 1406 TCGAGCTGGTGTGCTGACATCCGCCTGCGCGTACCGCG 1347
 Qy 303 AACTGGCGGACAGGGAAATGTGCGTAGTGTACTGGCTGACAGAAGTCG 302
 Db 1346 AGCTGGCGGTGCGCTCGAGGGGGATCATCTGTATCACCGGGCGAACAGGATCG 1287
 Qy 363 ATAAATTCTCGGCCCTGAAATGCGCAGTGACTACATACACCAACCGTTAACCGC 422
 Db 1286 ATGCCATGCTGGCTGGAGTGGGGGACATCTGTATCACCAACCGTTAACCGC 422
 Qy 423 GTGAACGTGACATTGTGACGGCAACCTACTGTCCTGACATCTGGTACTGTCA 482
 Db 1169 GCGCCGCCCCCGCCGGCAGGCTCAGGGAGTCTGGGACTGGCTGCG 1110
 Qy 443 GCGAACGTGCTGCGGGC--GAAGATCTGATCGCCGGGGTGCCTACCCAGGCC 499
 Db 1226 GCGAACGTGCTGCGGGC--GAAGATCTGATCGCCGGGGTGCCTACCCAGGCC 499
 Qy 483 GCGAACGTGCTGCGGGCAGGTTACAGTCTGATGGTGGAAACTACTGTCTGGTACCATGTA 542
 Db 1109 GGCGCCGCTGATGACCCACGGGAAACCTGAGTCCTGCTAACCCATGGGAAGA 602
 Qy 500 GCGCCGCCCCGGCCGGCAGGCTCAGGGAGTCTGGGACTGGCTGCG 559
 Db 1049 TGCTCCGGCTTCTCGCGACAGGGGATACTGTGACCGGAGTTCGGG 602
 Qy 543 GCGAACGTGCTGCGGGCAGGTTACAGTCTGATGGTGGAAACTACTGTCTGGTACCATGTA 542
 Db 1119 GGCGCCGCTGATGACCCACGGGAAACCTGAGTCCTGCTAACCCATGGGAAGA 602
 Qy 559 GCGAACGTGCTGCGGGCAGGTTACAGTCTGATGGTGGAAACTACTGTCTGGTACCATGTA 542
 Db 1109 GGCGCCGCTGATGACCCACGGGAAACCTGAGTCCTGCTAACCCATGGGAAGA 602
 Qy 603 CCATGCTACTTCTGGAACACGGCAATTGCGTGGCTGAGCTGCG 619
 Db 989 AGATCCGCAACCGCAGTGTGCGCTGCG 930
 Qy 620 TGCCTGGCGTTCTGCGAACACGGGATACTCTGACCCGGCACAGTGTGAGACC 679
 Db 929 TGCCTGAACATTTGGAATCTACGGGGATAACGGGAAATCATCCACCATTCAGGGTG 782
 Qy 633 AAATGACCGCCGTTGAGCTGAACACGGCACACCGTACTGTAGCTGACGATCCGGTA 722
 Db 929 TGCCTGAACATTTGGAATCTACGGGGATAACGGGAAATCATCCACCATTCAGGGTG 782
 Qy 660 GCGCCGCCCCGGCCGGCAGGCTCAGGGAGTCTGGGACTGGCTGCG 619
 Db 929 TGCCTGAACATTTGGAATCTACGGGGATAACGGGAAATCATCCACCATTCAGGGTG 782
 Qy 680 AGATCCGCAACCGCAGTGTGCGCTGCG 930
 Db 929 TGCCTGAACATTTGGAATCTACGGGGATAACGGGAAATCATCCACCATTCAGGGTG 782
 Qy 723 TTGGTAAACATTTGGAATCTACGGGGATAACGGGAAATCATCCACCATTCAGGGTG 782
 Db 869 CCGCTTACCTGTTACCG 852
 Qy 740 TGCCTGAACATTTGGAATCTACGGGGATAACGGGAAATCATCCACCATTCAGGGTG 782
 Db 800 CGCGCTTACCTGTTACCG 817
 RESULT 8
 US-09-252-991A-16309
 Sequence 16309, Application US/09252991A
 ; Paten No. 655175
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO *PSEUDOMONAS*
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIORITY APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIORITY APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO: 16309
 ; LENGTH: 3675
 ; TYPE: DNA
 ; ORGANISM: *Pseudomonas aeruginosa*
 US-09-252-991A-16309